Infrared behaviour and renormalization scheme invariance of QCD observables

Brooks, Paul M.

Infrared behaviour and renormalization scheme invariance of QCD observables

A thesis presented for the degree of

Doctor of Philosophy

by

Paul M. Brooks

Institute for Particle Physics Phenomenology

University of Durham

September 2006

Abstract

In this thesis we study the infrared (IR) behaviour of QCD observables, and solutions to the problem of renormalization scheme dependence. We investigate the IR behaviour of all-orders leading-b renormalon resummations for certain Euclidean observables (the Adler-D function and the GLS and Bjorken sum rules) in the Borel representation. We find that these resummations are finite at the Landau pole ($Q^2 = \Lambda^2$) and also 'freeze' to zero in the $Q^2 \to 0$ limit. We find this finite Landau pole behaviour has its origin in curious relations between IR and UV renormalons, which correspond to deeper conformal symmetries in QCD Green's functions. We consider these Borel resummed results in a skeleton expansion representation. This representation leads naturally to the standard Borel representation in the UV ($Q^2 > \Lambda^2$) region and to a modified Borel representation in the IR ($Q^2 < \Lambda^2$) region. We also consider the ambiguous part of the perturbative expansion in these representations. By demanding that such ambiguities cancel with similar ambiguities generated by the non-perturbative OPE, we are led to a new model for power corrections. We apply the complete renormalization group improved (CORGI) approach to all-orders renormalon resummations of the above-mentioned sum rules and compare the resultant predictions with experimental data. We also test our model for power corrections on these observables and find that the data favours power corrections of reasonably small magnitude.

We also apply the CORGI approach, together with the physical scale and the effective charge approaches, to moments of $F_3^{\nu N}$ and F_2^{ep}. We use the Bernstein averages method in which any dependence of the analysis on regions of x and Q^2 inaccessible to experiment is reduced. We also make use of the recently completed calculation for the NNLO anomalous dimension coefficients. We find that data for $F_3^{\nu N}$ favours a value of $\Lambda_{\overline{\rm MS}}^{(5)} = 219.1^{+23.6}_{-22.1}$ MeV, corresponding to $\alpha_s(M_Z) = 0.1189 \pm 0.0019$. In the case of F_2^{ep} we find $\Lambda_{\overline{\rm MS}}^{(5)} = 267.3^{+23.4}_{-19.1}$ MeV, corresponding to $\alpha_s(M_Z) = 0.1226^{+0.0012}_{-0.0017}$.

Acknowledgements

First and foremost, I would like to say thanks to my supervisor, Chris Maxwell, for all the advice and guidance he has provided.

Thanks also to PPARC for providing me with a studentship to fund this research.

Special thanks must go to the best office mates one could wish for. To Angelique for proving that it *is* possible to finish a PhD in three years. To Mark, for all his excellent guitar advice. To Rich (and Sarah) for their excellent BBQs, for Halo mayhem and for being such great friends. Finally to Tom, for crossword fun, being as inept on the guitar as I am and also for looking more nervous on my wedding day than I did. Thanks must also go to my second set of office mates, Gareth, Richard, Simon and Steve for putting up with me during the latter stages of my PhD.

I would also like to express my gratitude to Andrei Kataev, for his invaluable comments on my work on sum rules, renormalons and structure functions. Thanks also to Irinel Caprini, Jan Fischer, and Dimitri Shirkov for useful comments on the IR freezing paper, and to Andreas Vogt for useful discussions on quark mass thresholds.

Thanks also to my team of proof readers, Patricia, Tom and Julie, for pointing out all the obvious (and not so obvious) errors in my thesis. The write-up would have taken a lot, lot longer without you. Thanks also for not laughing too much at my comedy typos.

I would also like to say thank you to my parents, who have always been there for me.

Finally, to Julie, my wife. For the seemingly unending amounts of love, support and encouragement she has given me. I could not have done this without you.

Declaration

I declare that no material presented in this thesis has been previously submitted for a degree at this or any other university. No material from it should be published in any form without the author's prior written consent. All information derived from this thesis must be acknowledged appropriately.

The research described in this thesis has been carried out in collaboration with Dr. C. J. Maxwell. Chapters 4, 5 and 6 are based on the following publications:

- 'Infrared freezing of Euclidean QCD observables'
  P. M. Brooks and C. J. Maxwell
  Phys. Rev. **D74**, 065012 (2006) [arXiv:hep-ph/0604267]
- 'Comparison of NNLO and all-orders estimates of corrections to the GLS and Bjorken sum rules'
  P. M. Brooks and C. J. Maxwell
  [arXiv:hep-ph/0608339]
- 'Improved analysis of moments of F_3 in neutrino-nucleon scattering using the Bernstein polynomial method'
  P. M. Brooks and C. J. Maxwell
  [arXiv:hep-ph/0610137]

The research carried out in chapter 7 is currently in preparation for publication.

Contents

List of Figures

List of Tables

Preface

Over the last three decades, quantum chromodynamics (QCD) has firmly established itself as the fundamental theory of strong interactions. However, despite its innumerable successes, there remain several major obstacles which limit the precision with which we can predict observables. In this thesis, we will investigate two of these: the infrared (low-energy) behaviour of observables, and the problem of renormalization scheme (RS) dependence.

In perturbative QCD, observables are approximated by an expansion in powers of the strong coupling constant. Truncating this expansion at some finite order forms the basis of fixed-order perturbation theory and it is assumed that the convergence properties of the perturbative expansion are such that this will provide an accurate approximation to the exact result. However, the energy dependence of the coupling induced by the renormalization process, combined with the non-Abelian SU(3) nature of QCD, leads to the coupling becoming large at low energies. This seriously compromises the convergence properties of the perturbative expansion, limiting the accuracy with which we can apply perturbation theory. This limit in accuracy can be quantified as a theoretical error, the magnitude of which is determined by estimating the size of the omitted higher order corrections. At the Landau pole the coupling becomes singular, representing a complete breakdown of perturbation theory and as a result, no meaningful predictions can be obtained at energies below this point.

This is compounded by the fact that the perturbative expansions for all QCD observables are known to be factorially divergent, with the n^{th}-order coefficients behaving as $n!$ for large n. Consequently, the perturbative approximation can only ever be asymptotic to the exact form of the observable, and therefore there is an inherent theoretical

ambiguity associated with perturbation theory. This ambiguity is present even at high energies and grows in magnitude as the coupling becomes larger. Furthermore, it is present no matter how many terms in the perturbation series we calculate and can only be negated when we include the full set of perturbative and non-perturbative contributions.

The disastrous Landau-pole behaviour of the coupling is related to the fact that in fixed-order perturbation theory we ignore an infinite set of terms present in the full expansion. Presumably when we sum over the full set of terms, this singular behaviour disappears, yielding a result which has more meaningful infrared (IR) behaviour. Fortunately, we have access to results for the leading-N_f component of the perturbative coefficients at all-orders for many observables. These results can be converted into the so-called leading-b form and resummed such that we obtain expressions which represent a significant subset of the terms present in the full perturbative expansion, summed to all-orders. The possibility that such resummations might yield well-defined results at or below the Landau pole, and hence improve the predictive power of perturbative QCD at low energies, forms a significant part of the research carried out in this thesis.

As mentioned above, the theoretical ambiguity associated with the asymptotic nature of perturbation theory can only be removed when we take non-perturbative effects into account; both non-perturbative and perturbative sectors generate such ambiguities, and in the full theory these compensate each other. This implies a direct correspondence between perturbative and non-perturbative physics in QCD, and represents a unique bridge between these two sectors. In contrast to the success of perturbative methods in QCD, non-perturbative effects remain poorly understood and any means by which we might gain insight into them would be very welcome. Since the nature of the perturbative ambiguities only becomes apparent when we consider the large-order behaviour of perturbation theory, the all-orders resummations mentioned above are an appropriate tool with which to study this correspondence. In doing so, it is hoped that we can nullify the theoretical ambiguity associated with the perturbative part, and furthermore, gain insight into the nature of the non-perturbative effects.

We will also investigate potential solutions to the problem of RS dependence. The renormalization procedure allows us to obtain predictions for observables which are well-defined and finite order-by-order in perturbation theory. However, the price we

pay for this is a residual dependence of the predictions on an arbitrary scale (the renormalization scale), and on the scheme in which we choose to perform the renormalization. Although the full perturbative expansion must be and *is* RS invariant, truncation at finite order yields a result which has the RS dependence which would otherwise have been cancelled by the omitted higher-order terms. Consequently, we have an additional source of theoretical error in our predictions. The standard approach to dealing with this problem is to set the renormalization scale equal to some energy scale with physical meaning, such as the centre of mass energy, and then choose a popular or convenient RS such as modified minimal subtraction. However, there is no compelling physical reason for either of these choices nor is there a consensus amongst physicists. Hence we are left with a prediction which is inherently ambiguous due to the freedom of choice in RS.

Various proposed solutions to this problem have been suggested and amongst the candidates is the complete renormalization group improved (CORGI) approach to perturbation theory. In this approach, the necessity of choosing a renormalization scale is removed by relating the centre of mass energy to the renormalization scale in an RS invariant manner. A final result can then be obtained which is invariant under changes in renormalization scheme or scale. For observables which require factorization it is necessary to choose both an RS *and* a factorization scheme (FS) and this results in a double dose of the theoretical ambiguity associated with RS dependence. Fortunately, the CORGI approach can be extended to treat both factorization and renormalization scheme (FRS) dependence. In this thesis we intend to test the accuracy of CORGI perturbation theory by applying it to a phenomenological analysis of sum rules and moments of structure functions in deep inelastic scattering.

Thesis outline

In chapter 1 we present an introduction to QCD. Particular attention is paid to the subject of renormalization scheme dependence; we describe some of the proposed solutions to this problem and how they can be applied to fixed-order perturbative predictions.

In chapter 2 we review the subject of deep inelastic lepton-hadron scattering. We define the moments of structure functions and show how their scale dependence can be derived. We also discuss the FS dependence of the moments, drawing parallels between

this and RS dependence.

Chapter 3 is intended as an introduction to the behaviour of perturbative QCD at large orders of the coupling. We discuss instantons and renormalons as sources of factorially divergent contributions to the perturbative coefficients. We then describe how the existence of renormalons has inspired a series of all-orders leading-N_f calculations and how these results can be given added significance by converting them into leading-b results. We also describe how the presence of renormalons generated by the IR behaviour of QCD causes the perturbative definitions of observables to be ambiguous, and how only when we include non-perturbative effects can we recover a self-consistent perturbative definition of observables.

In chapter 4 we take the renormalon inspired all-orders leading-b results for the Adler D-function and the GLS, polarized Bjorken and unpolarized Bjorken sum rules and study their behaviour at low energies. Specifically we are interested in their behaviour at the Landau pole, and whether they 'freeze' to a well-defined finite value as $Q^2 \to 0$. We also study the perturbative ambiguities generated by these resummations and investigate whether we can use them to infer any properties of the non-perturbative contributions to observables.

Chapter 5 is a phenomenological study of the above-mentioned sum rules, based on the findings of chapter 4. We adopt the CORGI approach to perturbation theory and adapt it to the all-orders leading-b resummations. We compare the resultant predictions with the available experimental data and also with the equivalent fixed-order perturbative predictions.

In chapter 6 we turn our attention to the FRS dependence of fixed-order predictions for moments of the F_3 structure function. We consider three different approaches to dealing with this problem: the CORGI, physical scale and effective charge approaches. We apply these approaches to a phenomenological analysis of the moments via the method of Bernstein averages, our intent being to obtain a prediction of the QCD scale parameter, Λ.

In chapter 7 we present an analysis similar to that carried out in chapter 6 but for moments of the F_2 structure function. The principal difference between the F_3 and

F_2 analyses is the existence of a singlet contribution to the scale dependence of the moments in the F_2 case. We apply the CORGI and physical scale approaches and attempt to extract a value of the QCD scale parameter from F_2 data.

Chapter 1

Perturbative QCD

The purpose of this thesis is to improve both the predictive power and our understanding of quantum chromodynamics (QCD). As a first step, it is necessary to understand the basics of how QCD forms a consistent field theory from which accurate predictions for physical observables can be derived. To this end, we review in this chapter the theoretical framework of perturbative quantum field theory, applied to QCD. More detailed and thorough reviews of QCD can be found in the following references [1–13].

1.1 Introduction

We are interested in the dynamics of particles – in broad terms, how they react when they encounter or are 'scattered-off' each other. Ideally we would like to calculate cross sections, as these are the most important physical quantities in this respect. This chapter describes the link between the theoretical content of QCD (characterized by the Lagrangian and its field content) and predictions for physical observables. Accordingly, the material is organized as follows:

We define the QCD Lagrangian and describe the central role it plays in defining the theory of strong interactions. In order to derive predictions for physical observables from the QCD Lagrangian, it is necessary to take the intermediate step of calculating Green's functions. From these it is reasonably easy to extract cross sections and therefore we concentrate mainly on how quantized Green's functions may be obtained, through the technique of path integral quantization. Next, we describe how the QCD

Feynman rules are obtained before proceeding to explain how QCD must be systematically renormalized in order to remove the singularities that are encountered when one tries to make non-trivial calculations. We then describe how the renormalization procedure generates an energy dependence for the strong coupling constant and how from this, we can infer the properties of asymptotic freedom and confinement.

A large proportion of this thesis is concerned with the renormalization scheme (RS) dependence of QCD observables. Therefore, we describe the source of this problem and detail three proposed approaches to overcoming it: the physical scale (PS), the complete renormalization group improved (CORGI) and the effective charge (EC) approaches. We emphasize how the RS-dependence of a calculation can always be parameterized, and how as a result of this, a set of RS invariant parameters appears naturally.

We briefly describe the nature of non-perturbative corrections to QCD observables and in so doing, introduce the operator product expansion. Finally, we define some of the observables that are of interest to us in this thesis.

1.2 The QCD Lagrangian

Our starting point is classical field theory and the definition of the action in four dimensions,

$$S = \int \mathcal{L}\, d^4x. \tag{1.1}$$

$\mathcal{L}$ is the four dimensional Lagrangian density (hereafter referred to simply as the Lagrangian) and is a function of fields and their derivatives. The Lagrangian serves a dual purpose. It is both a formal mathematical object from which we can derive predictions for physical observables, and also a compact means by which we may express the principal symmetries and field content of a theory.

The conjecture that QCD possesses SU(3) gauge symmetry leads us to propose that the following classical Lagrangian,

$$\mathcal{L}_{\rm cl} = \overline{\psi}_i(i\not{D} - m)_{ij}\psi_j - \frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a, \tag{1.2}$$

describes strong interactions.

The field content of $\mathcal{L}_{\rm cl}$ is as follows: ψ and $\overline{\psi}$ denote fermion (quark) fields, they are colour triplets as denoted by the indices i and j. There are six known flavours of fermion, hence the first term in Eq. (1.2) must be summed over all six of these, with a different mass, m_f, for each flavour f. This sum, together with all spinorial indices, has been suppressed for the sake of brevity. $F^a_{\mu\nu}$ is the gluon field strength tensor and is defined as,

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g f^{abc} G^b_\mu G^c_\nu. \tag{1.3}$$

The G^a_μ symbol denotes a set of vector boson (gluon) fields; their eight colour degrees of freedom are denoted by the indices a, b and c.

The covariant derivative, D^μ, is a generalized form of the partial derivative ∂^μ and is defined such that $\mathcal{L}_{\rm cl}$ is invariant under SU(3) transformations of the fields. In Eq. (1.2) the covariant derivative appears with a 'slash' through it. This denotes that it has been contracted with a gamma matrix, γ_μ. The defining property of gamma matrices is that they obey the following anti-commutation relation,

$$\{\gamma_\mu, \gamma_\nu\} = 2 g_{\mu\nu}, \tag{1.4}$$

where $g_{\mu\nu}$ is the space-time metric. When acting on quark and gluon fields, D_μ has the following form,

$$D^\mu_{ij}\psi_j = (\partial^\mu \delta_{ij} + ig[T \cdot G^\mu]_{ij})\psi_j, \tag{1.5}$$

$$D^\mu_{ab} G^\mu_b = \partial^\mu G^\mu_a + g f_{abc} G^\mu_b G^\nu_c. \tag{1.6}$$

Here, T_a are matrices of the fundamental representation of SU(3), and we have also introduced the notation $T \cdot G_\mu = T_a G^a_\mu$. These matrices are defined in terms of the eight Gell-Mann matrices, λ_a,

$$T_a = \frac{\lambda_a}{2}. \tag{1.7}$$

f_{abc} are the structure constants of SU(3), they are related to T_a by,

$$[T_a, T_b] = i f_{abc} T_c. \tag{1.8}$$

Expressions for λ_a can be found in many textbooks. Several useful identities involving the matrices T_a, are given in appendix A.

The Lagrangian of Eq. (1.2) possesses SU(3) local gauge symmetry, by which we mean that the following transformations,

$$\psi_i \longrightarrow \Omega_{ij}(x)\psi_j = \exp\{igT \cdot \Lambda(x)\}_{ij}\psi_j, \tag{1.9}$$

$$T \cdot G_\mu \longrightarrow \Omega(x)\left[T \cdot G_\mu + \Omega^{-1}(x)\frac{i}{g}\partial_\mu\Omega(x)\right]\Omega^{-1}(x), \tag{1.10}$$

leave $\mathcal{L}_{\rm cl}$ unchanged.[1] From $\mathcal{L}_{\rm cl}$ the classical QCD equations of motion can be derived, via the Euler-Lagrange equations:

$$(i\not{D} - m)\psi = 0, \tag{1.11}$$

$$D_\mu F_a^{\mu\nu} = g\overline{\psi}\gamma^\mu T_a \psi. \tag{1.12}$$

We recognize Eq. (1.11) as the Dirac equation and together, these equations describe the classical theory of strong interactions.

Gauge fixing

It remains for us to quantize our theory. However, if we were to attempt to do so at this stage, it would become apparent that quantizing $\mathcal{L}_{\rm cl}$ in a meaningful way is impossible, for the following reason. In order to maintain manifest Lorentz invariance, $\mathcal{L}_{\rm cl}$ is written purely in terms of four-vectors. However, for this we pay a high price. The (massless) gluon, like the photon, has only two physical degrees of freedom and yet the gluon field variable G_μ has four. As a result of this inconsistency, were we to attempt to form a perturbation series using $\mathcal{L}_{\rm cl}$ we would end up integrating over a set of field configurations that are related by the gauge transformations of Eqs. (1.9) and (1.10). This procedure (equivalent to carrying out the integral $\int_0^\infty dx$) inevitably results in the appearance of singularities when we attempt to perform even the most trivial of calculations, such as defining the gluon propagator. The most common approach to remedying this problem is to factor out this set of field configurations. This can be achieved through the addition to the Lagrangian of a 'gauge fixing' term, $\mathcal{L}_{\rm gf}$. Among the many possible forms $\mathcal{L}_{\rm gf}$ may take, one popular choice – the 'covariant gauge' – has

[1] From this point forward, we suppress i, j indices.

the form,

$$\mathcal{L}_{\rm gf} = -\frac{1}{2\xi}(\partial^\mu G^a_\mu)^2, \tag{1.13}$$

where ξ is known as the gauge parameter. Details of non-covariant gauges can be found in Refs. [14,15]. The inclusion of a gauge fixing term and the choosing of a value for the gauge parameter breaks the manifest gauge invariance of the Lagrangian. In order to restore this symmetry, we must ensure that all physical predictions are independent of the choice of gauge and therefore independent of ξ. On the rare occasion in this thesis when we have to choose a gauge, we set $\xi = 1$, which corresponds to the Feynman gauge.

The procedure of gauge fixing is necessary in both Abelian *and* non-Abelian field theories. However, in non-Abelian theories $\mathcal{L}_{\rm gf}$ must be supplemented by another term, $\mathcal{L}_{\rm fp}$, known as the Fadeev-Popov or 'ghost' Lagrangian [16]. It has the following form,

$$\mathcal{L}_{\rm fp} = \partial_\mu \eta^\dagger_a(\partial^\mu \eta_a + g f_{abc}\eta_b G^\mu_c). \tag{1.14}$$

Unitarity of the S-matrix is then ensured by BRST symmetry [17,18]. We have introduced a new set of fields, η_a, known as ghost fields. They are unphysical in the sense that they only propagate as intermediate states, undetectable in the initial and final states of any interaction. They do however appear in the QCD Feynman rules, acting to cancel the unphysical gluon polarization states. Therefore, ghost field contributions must always be included in any calculation.[2]

When we include gauge fixing and ghost terms, the final QCD Lagrangian has the form,

$$\mathcal{L}_{\rm QCD} = \mathcal{L}_{\rm cl} + \mathcal{L}_{\rm gf} + \mathcal{L}_{\rm fp}. \tag{1.15}$$

There are additional possible terms which could be added to $\mathcal{L}_{\rm QCD}$, all maintaining gauge invariance. However, these are not of concern in this thesis and we shall restrict ourselves to the minimal Lagrangian of Eq. (1.15), and consider this to be the fundamental object defining QCD.

[2]There *do* exist certain choices for the gauge fixing term which make the addition of a ghost term unnecessary. These are known as 'physical gauges'.

The QCD coupling constant

So far we have neglected to identify the symbol g, first introduced in Eqs. (1.3), (1.5) and (1.6). It appears wherever trilinear or quadrilinear products of fields occur and we therefore associate it with the coupling together of fields i.e. interactions. Known as the coupling constant, g is (in four dimensions) a dimensionless parameter and is characteristic of the strength of QCD interactions. The coupling constant is perhaps the most important parameter in QCD and from its nature and behaviour, we can gain much physical insight.

In light of this, it is often convenient to split $\mathcal{L}_{\rm QCD}$ into an interaction part and a 'free' part:

$$\mathcal{L}_{\rm QCD} \;=\; \mathcal{L}_0 + \mathcal{L}_I. \tag{1.16}$$

Here, $\mathcal{L}_0 \equiv \mathcal{L}(g=0)$, and consequently the interaction Lagrangian, $\mathcal{L}_I$, contains all the terms in $\mathcal{L}_{\rm QCD}$ which involve factors of g. QCD Green's functions, in the absence of interactions, are almost trivial and can be solved exactly. However, as soon as we 'turn' the interaction on (i.e. set $g \neq 0$), the resultant Green's functions become incalculable and it becomes necessary to use some approximation to extract information about them.

In the perturbative approximation we assume that $g \ll 1$, i.e. all interactions are 'weak'. The full Green's functions are then expanded in terms of the solutions to the free theory. A generic result of a perturbative calculation for an observable is an expansion in powers of g,

$$\mathcal{F}(g) \;=\; \sum_{n=0}^{\infty} f_n g^n. \tag{1.17}$$

The full result, $\mathcal{F}(g)$, can then be obtained by calculating the infinite set of coefficients, f_n. In practice, it is possible to calculate only a finite number of coefficients, *or*, in some cases we can calculate a portion of f_n for all n. The accuracy of the approximation rests on the assumptions that the above series converges *and* that f_n can be uniquely defined. Challenges to these assumptions form the central theme of this thesis.

Having discussed the basic ingredients of QCD, we now turn our attention to describing

how QCD becomes a true quantum theory, through the quantization procedure.

1.3 Path integral quantization

There exist two widely used methods of quantizing QCD: the canonical commutator method and the path integral method. In this thesis, we use the path integral quantization method, originally developed by Feynman and Dirac as an alternative means by which quantum mechanics may be formulated [19, 20]. Path integral quantization involves the use of functional integrals to relate the dynamical content of a theory (represented by the Lagrangian) to Green's functions. The method generalizes to the quantization of classical field theory, and provides an elegant method of formulating QFT. This section provides a brief summary of the main ideas behind this formalism, and shows how some of the important results relevant to later chapters are derived. A detailed discussion of path integral quantization can be found in many QFT texts, including the following, [1, 2, 6, 11–13].

The determination of Green's functions is central to the process of extracting physical information (in the form of observables) from a theory. In this sense, they are an intermediate step between the observables and $\mathcal{L}_{\rm QCD}$. Cross sections are closely related to transition amplitudes; these in turn can be related to Green's functions via the Lehmann-Symanzik-Zimmerman reduction formula (see Ref. [11] for further details of this). The step that is most difficult, and that also gives us greatest physical insight, is that of obtaining Green's functions from $\mathcal{L}_{\rm QCD}$, and this can be achieved through the path integral formalism.

1.3.1 The generating functional

We start by considering the simple example of the quantum field theory of a scalar field with a quartic interaction,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}\mu^2\phi^2 - \frac{1}{4!}\lambda\phi^4 \tag{1.18}$$

$$= \mathcal{L}_0 + \lambda\mathcal{L}_I, \tag{1.19}$$

where λ is the coupling in this theory. Subsequently, we will generalize the results obtained in this example and apply them to QCD.

It is necessary to define an object known as the vacuum to vacuum transition amplitude in the presence of a source, $J(x)$,

$$W[J] = N \int \mathscr{D}\phi \exp\left\{ i \int dx^4 (\mathcal{L}(\phi, \partial_\mu \phi) + J(x)\phi(x)) \right\}. \tag{1.20}$$

This is also known as the generating functional. $\mathscr{D}\phi$ represents the functional integral over all field configurations, $\phi(x)$, which satisfy certain boundary conditions, and the normalization factor N is defined so that $W[J=0] = 1$. We simplify things by first considering free scalar field theory; in which $W[J] = W_0[J] \equiv W[J]|_{\lambda=0}$. In this case we can obtain an exact expression for $W_0[J]$,

$$W_0[J] = N \int \mathscr{D}\phi \exp\left\{ -\frac{1}{2} \int d^4x \int d^4y \, \phi(x) k(x,y) \phi(y) + i \int d^4x J(x)\phi(x) \right\} \tag{1.21}$$

$$= N \int \mathscr{D}\phi \exp\Big(-\frac{1}{2}(\phi, k\phi) + i(J,\phi) \Big), \tag{1.22}$$

where $k(x,y) = \delta(x-y)(\partial_\mu \partial^\mu + m^2)$ and the notation $(\phi, k\phi)$ is self-explanatory. When we include the normalization factor N, Eq.(1.22) becomes,

$$W_0[J] = \exp\Big\{ -\frac{1}{2}(J, k^{-1} J) \Big\}. \tag{1.23}$$

Using the form of $k(x,y)$ given above it can be shown that,

$$k^{-1}(x,y) = i\Delta_F(x-y) \equiv i \int \frac{d^4p}{(2\pi)^4} \, e^{-ip.(x-y)} \frac{1}{p^2 - m^2 + i\varepsilon}, \tag{1.24}$$

$$\Longrightarrow W_0[J] = \exp\left\{ -\frac{i}{2} \int d^4x \int d^4y \, J(x) \Delta_F(x-y) J(y) \right\}, \tag{1.25}$$

where $\Delta_F(x)$ is known as the Feynman propagator. The integral in (1.24) represents a Fourier transform and so the Feynman propagator in momentum space can be defined as,

$$\tilde{\Delta}_F(p) \equiv \frac{1}{p^2 - m^2 + i\varepsilon}. \tag{1.26}$$

Because our scattering experiments involve external states of definite momentum, we express Green's functions in terms of components defined in momentum space.

We have obtained an exact expression for $W[J]$ in the case of free scalar field theory. The path integral contains an oscillatory exponential and appears to be ill-defined.

However this can be remedied by performing a Wick rotation of the time coordinate, $t \rightarrow it$, evaluating $W_0[J]$ in Euclidean space and then continuing back to Minkowski space. This process is justified provided there exist no poles above the positive, real semi-axis or below the negative real semi-axis. The Feynman prescription, introduced via an $i\varepsilon$ term in Eq. (1.24), ensures that this is the case. The introduction of the source $J(x)$ can be understood, in broad terms, as the introduction of a probe field as a means by which we may examine the vacuum of the theory. The form of the generating functional encapsulates the result of that examination. When we come to consider QCD, things are slightly more complicated; it will become necessary to introduce source terms for each type of field.

1.3.2 Green's functions

The next step is to show how Green's functions are obtained from the generating functional. An n-particle Green's function is defined as the vacuum expectation value of n time ordered fields and is denoted by $\mathcal{G}^{(n)}(x_1, x_2, ..., x_n)$. This can be related to $W[J]$ by extending to field theory a result from the path integral quantization of QM [2]:

$$\begin{aligned} \mathcal{G}^{(n)}(x_1, x_2, ..., x_n) &= \langle 0|T(\phi(x_1)\phi(x_2)...\phi(x_n))|0\rangle \\ &= (i)^{-n} \frac{\delta^n W[J]}{\delta J(x_1)\delta J(x_2)...\delta J(x_n)}\bigg|_{J=0}. \end{aligned} \tag{1.27}$$

The Green's functions for *free* scalar field theory, though useful as a gentle introduction, are rather uninteresting; we will extend our analysis to include interactions before we investigate Eq. (1.27) any further.

When we turn the interactions on, Eq. (1.20) becomes

$$W[J] = N \int \mathscr{D}\phi \exp\left\{i\lambda \int d^4x\, \mathcal{L}_I(\phi)\right\} \exp\left\{i \int d^4x\, (\mathcal{L}_0 + J\phi)\right\} \tag{1.28}$$

$$= N \int \mathscr{D}\phi \left\{\sum_{n=0}^{\infty} \frac{(i\lambda)^n}{n!} \left[\int d^4x\, \mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right]^n\right\} \exp\left\{i \int d^4x\, (\mathcal{L}_0 + J\phi)\right\}. \tag{1.29}$$

In Eq. (1.29), the exponential of the interaction action has been expanded as a Taylor

series in powers of λ. Also, in $\mathcal{L}_I(\phi)$ each instance of ϕ is replaced by a functional derivative with respect to $J(x)$; this replacement builds the ϕ dependence of $\mathcal{L}_I(\phi)$. The purpose of this replacement is to make $\mathcal{L}_I$ independent of ϕ, allowing us to take the interaction piece outside of the integral,

$$\begin{aligned} W[J] &= N\left\{\sum_{n=0}^{\infty}\frac{(i\lambda)^n}{n!}\left[\int d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right]^n\right\}\int \mathscr{D}\phi\,\exp\left\{i\int d^4x\,(\mathcal{L}_0+J\phi)\right\} \\ &= N\,\exp\left[i\lambda\int d^4x\,\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x)}\right)\right]W_0[J]. \end{aligned} \tag{1.30}$$

Note also that integral and summation symbols have been interchanged. This is a subtle though significant step which can, in some cases, disrupt the convergence of perturbative expansions (see Refs. [21–23] for details of this). To make the perturbative application of Eq. (1.30) more explicit we rewrite it as,

$$W[J] = \sum_{n=0}^{\infty}\frac{(i\lambda)^n}{n!}\int d^4x_1...d^4x_n\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x_1)}\right)...\mathcal{L}_I\left(-i\frac{\delta}{\delta J(x_n)}\right)W_0[J]. \tag{1.31}$$

In summary, we have succeeded in writing the full generating functional as an expansion in powers of λ. Here we see the mechanism of the perturbative approximation in action. The exact, though incalculable, expression in Eq. (1.30) has been manipulated to produce an infinite series of calculable terms.

Using Eq. (1.31) in conjunction with Eq. (1.27), Green's functions can be expressed as an (infinite) series in powers of the coupling λ,

$$\mathcal{G}^{(n)}(\lambda) = \sum_{k=0}^{\infty}\mathcal{G}_k^{(n)}\lambda^k. \tag{1.32}$$

It then follows that a generic physical observable, $\mathcal{F}(\lambda)$, can also be expressed as a perturbative expansion,

$$\mathcal{F}(\lambda) = \sum_{n=0}^{\infty}f_n\lambda^n. \tag{1.33}$$

In this approximation, Green's functions can be determined by calculating terms in Eq. (1.32) at successive orders of λ, up to the highest order we have the strength to

calculate. The complete form of the Green's functions can be reconstructed by including an infinite number of terms; it is of course hoped that the series rapidly converges and that it is necessary to calculate only the first few terms. The convergence properties of Green's functions and physical observables are discussed in more detail in chapter 3.

Connected and one particle irreducible Green's functions

In practice it is often more productive to consider 'connected' Green's functions. These are derived from the connected generating functional, $\mathcal{X}[J]$, defined as,

$$\mathcal{X}[J] = -i \log W[J]. \tag{1.34}$$

The connected Green's functions, $X^{(n)}(x_1, x_2, \ldots, x_n)$, are then obtained from this generating functional using an expression analogous to Eq. (1.27). Furthermore, it is also beneficial to consider only a subset of connected Green's functions known as one particle irreducible (OPI) Green's functions, which we denote by $G^{(n)}$. The difference between normal connected and OPI Green's functions will become apparent when we come to consider their diagrammatic interpretation.

OPI diagrams can be derived from an object known as the effective action, $\Gamma[\phi_c]$ defined as a function of the classical field, ϕ_c,

$$\Gamma[\phi_c] = \mathcal{X}[J] - (J, \phi_c). \tag{1.35}$$

The connected OPI Green's functions are then defined in terms of the effective action; in analogy with Eq. (1.27),

$$G^{(n)}(x_1, x_2, ..., x_n) = (i)^{-n} \frac{\delta^n \Gamma[\phi_c]}{\delta\phi_c(x_1)\delta\phi_c(x_2)...\delta\phi_c(x_n)}\bigg|_{J=0}. \tag{1.36}$$

The effective action can also be used to illustrate how the quantization procedure has modified our theory. If we calculate $\Gamma[\phi_c]$ up to order λ, we obtain an expression in terms of integrals of classical fields, sources and propagators.[3] This can be solved to

[3] See Ref. [2] p. 66 for further details.

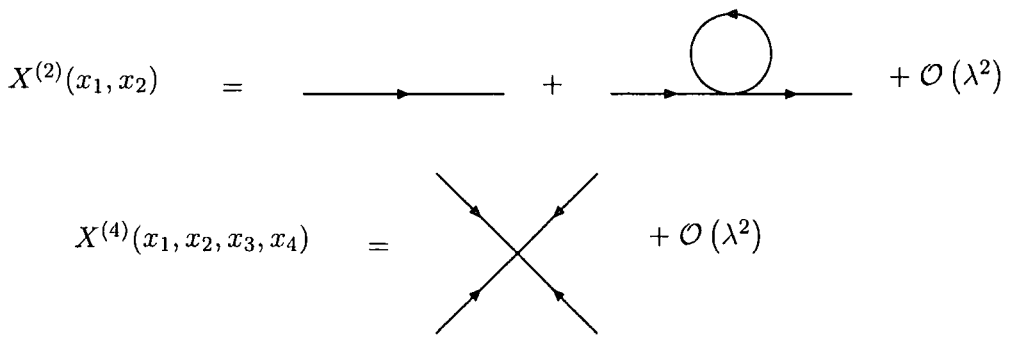


Figure 1.1: The 2- and 4-point connected Green's functions in ϕ^4, up to $\mathcal{O}(\lambda)$, given in Eqs. (1.39) and (1.40).

obtain an expression for $J(x)$ in terms of ϕ_c,

$$J(x) = \left(\Box + \mu^2 + \frac{1}{2}\lambda i\Delta_F(0)\right)\phi_c + \frac{1}{6}\lambda\phi_c^3 + \mathcal{O}(\lambda^2). \tag{1.37}$$

Setting $J(x) = 0$ (we can do this because it is an arbitrary function) gives an equation which we can consider to be the quantum equation of motion. This can be compared with the classical equation of motion for ϕ^4 theory, the Klein-Gordon equation:

$$\left(\Box + \mu^2\right)\phi + \frac{1}{6}\lambda\phi^3 = 0. \tag{1.38}$$

If we restored the factors of $\hbar$ in the Eq. (1.37), we would see that the $\Delta_F(0)$ term is proportional to $\hbar$. Hence this is a quantum correction, not present in the classical $\hbar \to 0$ limit, and generated by the quantization procedure.

For illustrative purposes, we now use Eqs. (1.27), (1.31) and (1.34) to calculate the 2- and 4-point connected Green's functions up to $\mathcal{O}(\lambda)$:

$$\begin{aligned} X^{(2)}(x_1, x_2) &= i\Delta_F(x_1 - x_2) - \frac{1}{2}i\lambda \int dx\, i\Delta_F(x_1 - x) i\Delta_F(x - x) i\Delta_F(x - x_2) \\ &+ \mathcal{O}(\lambda^2), \end{aligned} \tag{1.39}$$

$$\begin{aligned} X^{(4)}(x_1, x_2, x_3, x_4) &= -i\lambda \int dx\, \Delta_F(x_1 - x)\Delta_F(x_2 - x)\Delta_F(x_3 - x)\Delta_F(x_4 - x) \\ &+ \mathcal{O}(\lambda^2). \end{aligned} \tag{1.40}$$

The Feynman propagator in Eq. (1.24) is associated with the propagation of solutions to the Klein-Gordon equation. Hence $\Delta_F(x-y)$, and therefore the propagators in Eq. (1.39), may be represented diagrammatically as lines propagating between space-time points x and y. Equivalently, the momentum space propagator $\Delta_F(p)$ can be represented by a line carrying momentum p. Figure 1.1 shows such diagrammatic representations of the Green's functions in Eqs. (1.39) and (1.40). In this diagrammatic interpretation, connected diagrams are those which cannot be considered to be two separate diagrams; and OPI diagrams are diagrams that cannot be split into two separate diagrams by the removal of a single line.

It is clear that the Green's functions given in Eq. (1.39) and (1.40) and depicted in Fig. 1.1 can be written in terms of just two components: the Feynman propagator and a $-i\lambda$ term associated with each vertex. This is true of all n-point Green's functions at any order of perturbation theory. Consequently, we can obtain expressions of the form of Eq. (1.39) and (1.40) for any n-particle Green's functions without explicitly evaluating Eq. (1.27).

The Feynman rules provide a shorthand for deriving these Green's functions with as little work as possible. They consist of a set of mathematical expressions, in momentum space, one for each propagator and vertex in the Lagrangian, together with a set of rules for integrating over internal momenta, and necessary symmetry factors. These rules provide us with a systematic approach for calculating Green's functions order by order in perturbation theory [24].

Before we discuss the Feynman rules in more detail, we describe how the path integral quantization technique is adapted to QCD.

1.3.3 The QCD generating functional

The generating functional of QCD is slightly more complicated than that of ϕ^4 theory, given in Eq. (1.28). To derive it, we first split $\mathcal{L}_0$ into quark, gluon and ghost terms

i.e. terms bilinear in these fields or their derivatives,[4]

$$\mathcal{L}_0 = \mathcal{L}_0^q + \mathcal{L}_0^g + \mathcal{L}_0^{fp}. \tag{1.41}$$

Similarly, we split the $\mathcal{L}_I$ into quark-gluon three-gluon, four-gluon and ghost-gluon vertex terms,

$$\mathcal{L}_I = \mathcal{L}_I^{q\text{-}g} + \mathcal{L}_I^{3g} + \mathcal{L}_I^{4g} + \mathcal{L}_I^{fp\text{-}g}. \tag{1.42}$$

Next, we define the free generating functional, in analogy with Eq. (1.20),

$$W_0[J,\zeta,\zeta^\dagger,\chi,\overline{\chi}] = W_0^g[J]\, W_0^{fp}[\zeta,\zeta^\dagger]\, W_0^f[\chi,\overline{\chi}], \tag{1.43}$$

where,

$$W_0^g[J] = \int \mathscr{D}G \exp\left[i\int d^4x\,(\mathcal{L}_0^g + G\cdot J)\right], \tag{1.44}$$

$$W_0^{fp}[\zeta,\zeta^\dagger] = \int \mathscr{D}\eta\,\mathscr{D}\eta^\dagger \exp\left[i\int d^4x\,(\mathcal{L}_0^{fp} + \eta^\dagger\zeta + \zeta^\dagger\eta)\right], \tag{1.45}$$

$$W_0^q[\chi,\overline{\chi}] = \int \mathscr{D}\psi\,\mathscr{D}\overline{\psi} \exp\left[i\int d^4x\,(\mathcal{L}_0^q + \overline{\psi}\chi + \overline{\chi}\psi)\right], \tag{1.46}$$

and $J(x)$, $\zeta(x)$ and $\chi(x)$ are source terms associated with the three different types of field. Note that the fields $\psi(x)$, $\chi(x)$, $\eta(x)$ and $\zeta(x)$ are all Grassman variables and care must be taken to ensure they are treated as such upon integration and differentiation. In analogy with the scalar field theory example and Eq. (1.25), we can obtain exact expressions for these three functionals,

$$W_0^g[J] = \exp\left\{\frac{i}{2}\int d^4x\,d^4y\,J^{a\mu}(x)D^{ab}_{\mu\nu}(x-y)J^{b\nu}(y)\right\}, \tag{1.47}$$

$$W_0^{fp}[\zeta,\zeta^\dagger] = \exp\left\{i\int d^4x\,d^4y\,\zeta^a(x)^\dagger D^{ab}(x-y)\zeta^b(y)\right\}, \tag{1.48}$$

$$W_0^q[\chi,\overline{\chi}] = \exp\left\{i\int d^4x\,d^4y\,\overline{\chi}(x)S(x-y)\chi(y)\right\}, \tag{1.49}$$

[4]Note that $\mathcal{L}_0^g$ contains not only the Abelian terms in $-\frac{1}{4}F^a_{\mu\nu}F_a^{\mu\nu}$ but also the gauge fixing term, $\mathcal{L}_{\rm gf}$. Similarly $\mathcal{L}_0^{fp}$ is just the $\mathcal{O}(g^0)$ term in $\mathcal{L}_{\rm fp}$.

where $D(x)$ and $S(x)$ are the QCD field propagators in configuration space. Their Fourier transforms are given by,

$$\tilde{D}^{ab}_{\mu\nu}(p) = \frac{i\delta^{ab}}{p^2+i\varepsilon}\left(-g_{\mu\nu}+(1-\xi)\frac{p_\mu p_\nu}{p^2}\right), \tag{1.50}$$

$$\tilde{D}^{ab}(p) = \frac{i\delta^{ab}}{p^2+i\varepsilon}, \tag{1.51}$$

$$\tilde{S}_{ij}(p) = i\delta_{ij}\frac{\not{p}+m}{p^2-m^2+i\varepsilon}. \tag{1.52}$$

In analogy with Eq. (1.30), the full QCD generating functional can be written as,

$$W[J,\zeta,\zeta^\dagger,\chi,\overline{\chi}] = \exp\left[i\int d^4x\,\mathcal{L}_I\left(\frac{\delta}{i\delta J},\frac{-\delta}{i\delta\zeta},\frac{\delta}{i\delta\zeta^\dagger},\frac{-\delta}{i\delta\chi},\frac{\delta}{i\delta\overline{\chi}}\right)\right]W_0[J,\zeta,\zeta^\dagger,\chi,\overline{\chi}], \tag{1.53}$$

where each instance of each field in $\mathcal{L}_I$ has been replaced by a functional derivative with respect to the appropriate source term.

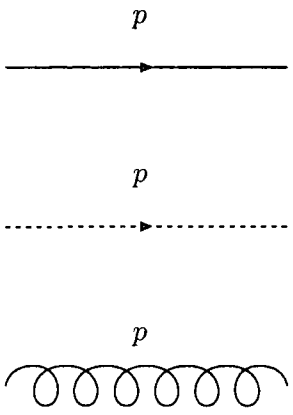
In analogy with Eq. (1.31) we can approximate Eq. (1.53) perturbatively to obtain an expansion in powers of g. Hence all QCD Green's functions can be obtained by using an equation analogous to Eq. (1.27):

$$\mathcal{G}^{(n)}(x_1,x_2,\ldots,x_n) = (i)^{-n}\left(\prod_i^n\frac{\delta}{\delta\mathcal{S}_i}\right)W[J,\zeta,\zeta^\dagger,\chi,\overline{\chi}], \tag{1.54}$$

where $\mathcal{S}_i$ can be either of the five source terms. The correct combination of these source term derivatives depends on the types of external field that the Green's function refers to.

1.4 QCD Feynman rules

We saw in section 1.3.2 how the propagators for ϕ^4 theory can be obtained by considering the free generating functional, isolating the term coupled to two sources and identifying it as the propagator (in configuration space). In general, field bilinear terms in the Lagrangian give rise to the propagators, and multilinear terms generate interaction terms in the Feynman rules. Hence, inspection of Eqs. (1.47) - (1.49) allows us to



$$= \quad i\delta_{ij}\frac{\not{p}+m}{p^2-m^2+i\varepsilon}$$

$$= \quad \frac{i\delta^{ab}}{p^2+i\varepsilon}$$

$$= \quad \frac{i\delta^{ab}}{p^2+i\varepsilon}\left(-g_{\mu\nu}+(1-\xi)\frac{p_\mu p_\nu}{p^2}\right)$$

Figure 1.2: Feynman rules for the propagators of QCD. Expressions for $\tilde{D}^{ab}_{\mu\nu}(p)$, $\tilde{D}^{ab}(p)$ and $\tilde{S}_{ij}(p)$ can be found in Eqs. (1.50) - (1.52).

obtain the propagators for QCD fields. These are given, along with their diagrammatic representations in Fig. 1.2.

To calculate the interaction Feynman rules, we isolate the term in the Lagrangian relevant to a particular vertex, calculate the generating functional for that vertex and then use Eq. (1.54) to obtain the relevant Green's function. For example, for the quark-gluon vertex, the relevant generating functional is,

$$W^{q\text{-}g}[J,\chi,\overline{\chi}] \;=\; i\int d^4x\,\mathcal{L}_I^{q\text{-}g}\left(\frac{\delta}{i\delta J^{a\mu}},\frac{-\delta}{i\delta\chi},\frac{\delta}{i\delta\overline{\chi}}\right)W_0^g[J]\,W_0^q[\chi,\overline{\chi}], \qquad (1.55)$$

and the relevant 3-point Green's function is,

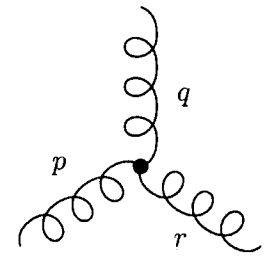
$$\mathcal{G}^{q\text{-}g,(3)}(x_1,x_2,x_3) \;=\; \frac{-\delta^3}{\delta\chi\,\delta\overline{\chi}\,\delta J}\,W^{q\text{-}g}[J,\chi,\overline{\chi}]. \qquad (1.56)$$

When converted to the momentum representation this gives,[5]

$$\tilde{\mathcal{G}}^{q\text{-}g,(3)}(p_1,p_2,k) \;=\; -igT^a\frac{i}{\not{p}_1-m}\gamma_\mu\frac{i}{\not{p}_2-m}\frac{i\,d_{\mu\nu}(k)}{k^2}, \qquad (1.57)$$

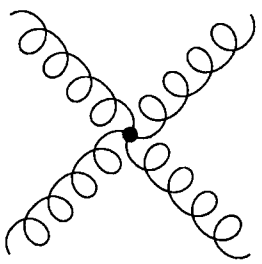
where p_1, p_2 and k are the momenta carried by the two quarks and the gluon, respectively and we have defined $d_{\mu\nu}\equiv -g_{\mu\nu}+(1-\xi)k_\mu k_\nu/k^2$. By stripping Eq. (1.57) of external propagators we can obtain the Feynman rule for the quark-gluon vertex. Its

[5]We have dropped the Lorentz and colour indices from $\mathcal{G}^{q\text{-}g,(3)}$ in order to make the expressions clearer.

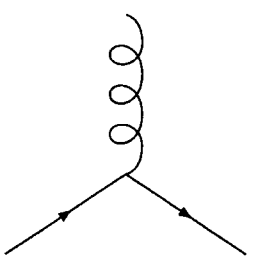


$$= \quad -g\, f^{abc}\left[(p-q)^{\sigma}\, g^{\mu\nu} + (q-r)^{\mu}\, g^{\nu\sigma} + (r-p)^{\nu}\, g^{\sigma\mu}\right],$$

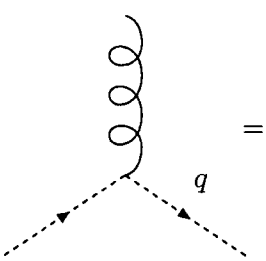
(all momenta incoming, $p+q+r=0$)



$$= \quad \begin{aligned} &-ig^2\, f^{xac} f^{xbd}\left[g^{\mu\nu} g^{\rho\sigma} - g^{\mu\sigma} g^{\nu\rho}\right] \\ &-ig^2\, f^{xad} f^{xbc}\left[g^{\mu\nu} g^{\rho\sigma} - g^{\mu\rho} g^{\nu\sigma}\right] \\ &-ig^2\, f^{xab} f^{xcd}\left[g^{\mu\rho} g^{\mu\sigma} - g^{\mu\sigma} g^{\nu\rho}\right], \end{aligned}$$



$$= \quad -igT^{a}\gamma^{\mu},$$



$$= \quad g f^{abc} q^{\mu}.$$

Figure 1.3: Feynman rules for the vertices of QCD.

form is given in Fig. 1.3, along with the rest of the QCD vertex Feynman rules.[6]

1.5 Renormalization

In section 1.3.2 we saw how the classical equations of motion for ϕ^4 theory are modified by the quantization procedure,

$$\left(\Box + \mu^2 + \frac{1}{2}\lambda i\Delta_F(0)\right)\phi_c \;=\; -\frac{1}{6}\lambda\phi_c^3 + \mathcal{O}(\lambda^2). \tag{1.58}$$

In a system of particles described by ϕ^4 theory, a measurement of particle mass would measure the terms in Eq. (1.58) that are proportional to ϕ_c; the measurement would

[6] See Ref. [1] for further details.

not be able to distinguish between the μ^2 and $\Delta_F(0)$ terms. If it were possible to turn the interactions off, we would be able to measure μ. However, this is impossible and we therefore conclude that μ is an unmeasurable (and therefore unphysical) quantity. Accordingly we redefine it as the 'bare' mass, μ_B. The measurable quantity would then be the effective or 'renormalized' mass, μ_R,

$$\mu_R^2 = \mu_B^2 + \frac{1}{2}\lambda i\Delta_F(0) + \mathcal{O}(\lambda^2) \tag{1.59}$$

$$\equiv Z\mu_B^2, \tag{1.60}$$

where Z is known as the mass renormalization constant. Were we to calculate higher-order terms in Eq. (1.58), it would become apparent that the same is true for the coupling, λ, and in fact for all parameters in the Lagrangian.

As a consequence of this, it would appear that all parameters in the Lagrangian of Eq. (1.18) are unphysical. Furthermore, the propagator term in Eq. (1.59) is singular, meaning that if we wish to recover a finite value for the measured mass, μ_R, then μ_B must also be singular. It must contain the same singularity as that in Eq. (1.59), but with the opposite sign, so that the singularities in Z and μ_B cancel each other out. The fact that μ_B is singular is troubling. However, we can justify this worrying turn of events by reminding ourselves that bare parameters are unphysical and hence they are allowed to be singular.

The singularity structure of Z can be determined by calculating the singular term in Eq. (1.59) and demanding that μ_R be finite. However, it must be noted that the value of Z can be shifted by an arbitrary finite amount, without violating this physicality condition. As we shall see later, this element of arbitrariness propagates into renormalized predictions for physical observables.

The type of divergence in Eq. (1.59) is endemic in field theory (in four dimensions). It occurs because field theory allows the existence of intermediate states which can propagate with arbitrarily large momenta. The same situation arises in QCD. Let us consider the order $\mathcal{O}(g^2)$ fermion loop correction to the gluon propagator, illustrated in Fig. 1.4. Using the Feynman rules in section 1.4 we can construct the following

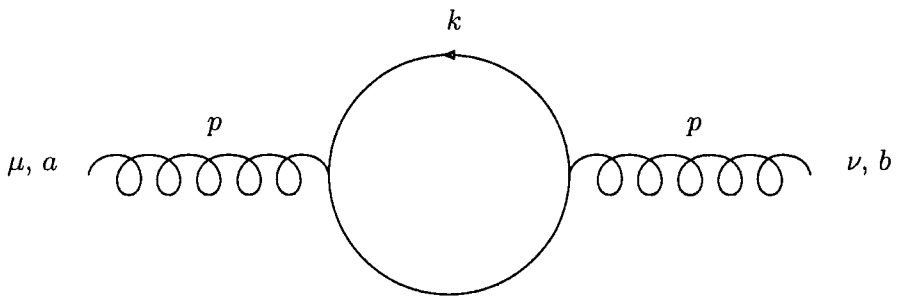


Figure 1.4: Fermion loop contribution to corrections to the gluon propagator.

expression,

$$\begin{aligned} \text{Figure 1.4} &\equiv -i\Pi^{\mu\nu}_{ab}(p^2) \\ &= -g^2\,\mathrm{Tr}(T_aT_b)\int_0^\infty \frac{d^4k}{(2\pi)^4}\frac{\gamma^\mu(\not{k}+m)\gamma^\nu(\not{k}+\not{p}+m)}{(k^2-m^2)([k+p]^2-m^2)}. \end{aligned} \tag{1.61}$$

Ignoring, for a moment, the numerator of the integrand and the colour structure, and using the following identity,

$$\frac{1}{xy} = \int_0^1 \frac{d\alpha}{(x(1-\alpha)+y\alpha)^2}, \tag{1.62}$$

we obtain,

$$\Pi^{\mu\nu}_{ab}(p^2) \propto \int_0^1\int_0^\infty \frac{d\alpha\, d^4k}{(k^2-M^2)^2}, \tag{1.63}$$

where the substitutions $k \to k-\alpha p$ and $M = m^2-\alpha(1-\alpha)p^2$ have been made.

The integral in Eq. (1.63) is logarithmically divergent, exhibiting the same kind of divergence present in Eq. (1.59). Fortunately, we have only encountered logarithmic divergences. Had the divergences been quadratic, the situation we find ourselves in would have been many times worse. However, in order to extract a meaningful result from Eq. (1.61) we must first regularize and then renormalize it.

1.5.1 Regularization

The purpose of regularization is to quantify the extent to which a particular diagram, such as Fig. 1.4, is divergent. This allows us to obtain expressions for Green's functions which, although still formally divergent, expose the nature of their singularities. A number of regularization procedures exist. However, we limit ourselves to the study of dimensional regularization (DR). First, we note that the divergent nature of $\Pi^{\mu\nu}_{ab}(p^2)$ is dependent on the number of space-time dimensions i.e. the number of k integrations. By generalizing (using analytic continuation) the number of dimensions to $D = 4 - 2\varepsilon$ we can control the divergent nature of Eq. (1.61); eventually we will set ε to zero in order to restore the $D = 4$ limit. In D dimensions, the integral in Eq. (1.63) can be obtained from the following result,

$$\int_0^\infty \frac{d^D k}{(2\pi)^D} \frac{1}{(k^2 - M^2)^{n+2}} = i\frac{(-1)^n}{(4\pi)^2} \frac{\Gamma(n+\varepsilon)}{\Gamma(n+2)} M^{-2n} \left(\frac{M^2}{4\pi}\right)^{-\varepsilon}. \tag{1.64}$$

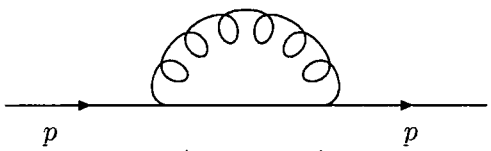
Generalizing the number of dimensions to D also has the effect of changing the mass dimension of the Lagrangian to D. Using this fact, and dimensional analysis, we can deduce that the fermion and gluon fields and the coupling constant, g, have mass dimensions of $3/2-\varepsilon$, $1-\varepsilon$ and ε respectively. In light of this, we redefine the (previously dimensionless) coupling g, in terms of a truly dimensionless coupling, $\bar{g}$, and an arbitrary scale μ,

$$g = \bar{g}\mu^\varepsilon. \tag{1.65}$$

The scale μ generates the mass dimension of g, and is known as the renormalization scale. Despite it being physically meaningless, the final result of any renormalized calculation to finite order in perturbation theory *will* be dependent on μ. This dependence on an arbitrary scale is a manifestation of the inherently ambiguous nature of regulated divergent integrals.

The numerator of the integrand in Eq. (1.61) is a trace over products of gamma matrices. Upon performing this trace, and after the application of the result in Eq. (1.64), one obtains,

$$\begin{aligned} -i\Pi^{\mu\nu}_{ab}(p^2) &= i\frac{\bar{g}^2}{16\pi^2}\frac{4T_R}{3}\delta_{ab}(p^\mu p^\nu - g^{\mu\nu}p^2)\left(-\frac{1}{\varepsilon} + \text{finite}\right) \\ &\equiv \delta_{ab}(p^\mu p^\nu - g^{\mu\nu}p^2)\Pi(p^2). \end{aligned} \tag{1.66}$$



(diagram 1)

Figure 1.5: The one-loop correction to the fermion propagator, also known as the self-energy diagram.

The presence of a $1/\varepsilon$ pole makes explicit the singular nature of this one-loop integral; it is a manifestation of the logarithmic divergence in Eq. (1.63). The finite part in Eq. (1.66) has the form,

$$\begin{aligned} \text{finite} &= \gamma_{\text{E}} + 6\int_0^1 d\alpha\, \alpha(1-\alpha)\ln\left(\frac{m^2-\alpha(1-\alpha)p^2}{4\pi\mu^2}\right) \\ &= \left[\gamma_{\text{E}} - \ln(4\pi) + \ln\left(\frac{-p^2}{\mu^2}\right) - \frac{5}{3}\right] - \frac{m^2}{p^2} + \mathcal{O}\left(\frac{m^6}{p^6}\right) + \mathcal{O}(\varepsilon). \end{aligned} \tag{1.67}$$

Where γ_{E} is the Euler-Mascheroni constant and has an approximate value of $\gamma_{\text{E}} = 0.577215665\ldots$. Terms of order ε or higher can be neglected (at one-loop) as they are zero in four dimensions. Furthermore, in this thesis we will be working in 'massless QCD' in which all quark masses are set to zero. Therefore we need consider only the terms in the square brackets in Eq. (1.67). The $1/\varepsilon$ pole in the vacuum polarization function in Eq. (1.66) is the crucial component of this result, important when we come to renormalize QCD. In general, higher-loop diagrams contain higher-order pole terms.

Figures 1.5 - 1.7 show 7 of the one-loop OPI diagrams that contribute to the quark and gluon propagators and the quark-gluon vertex. For illustrative purposes we shall now state their (dimensionally) regularized forms:

Diagram 1 is known as the fermion self-energy. Its regularized form is

$$(\text{diagram 1}) = i\frac{\overline{g}^2}{16\pi^2}\delta_{ij}C_F\left[\not{p} - 4m - (1-\xi)(\not{p}-m)\right]_{ab}\frac{1}{\varepsilon'} + \ldots, \tag{1.68}$$

where,

$$\frac{1}{\varepsilon'} = \frac{1}{\varepsilon} + \ln 4\pi - \gamma_{\text{E}} - \ln\frac{-p^2}{\mu^2}, \tag{1.69}$$

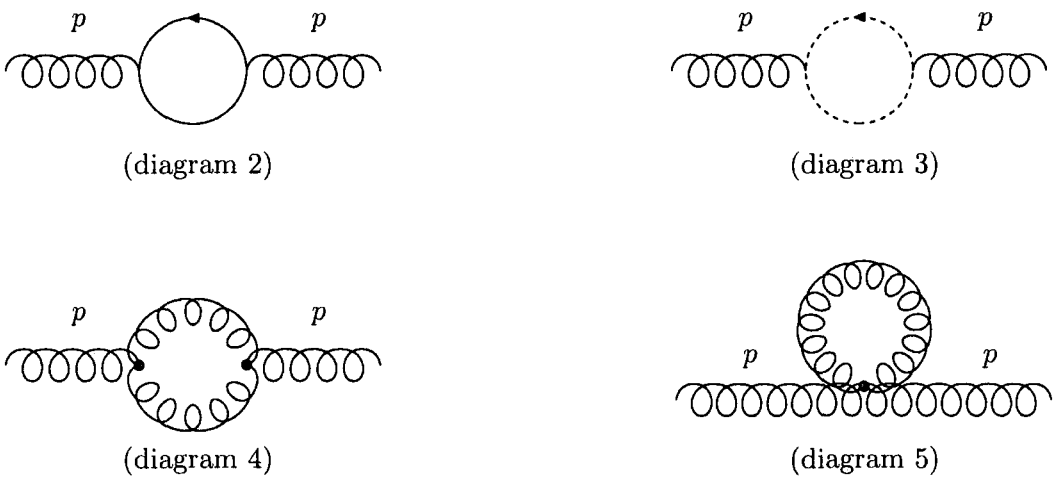


Figure 1.6: One-loop corrections to the gluon propagator, also known as vacuum polarization diagrams.

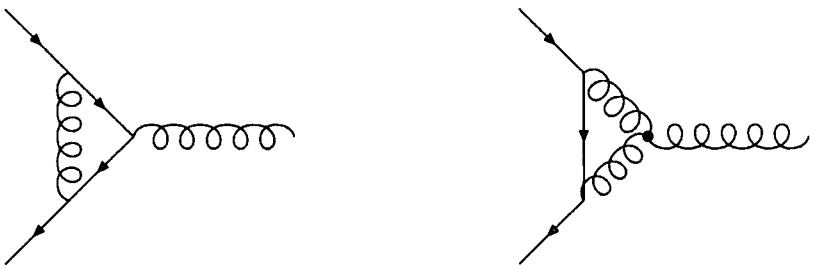
and the ellipsis represents terms absent or finite in the $\varepsilon \to 0$ limit. We have kept the m dependence but from now on, we will set $m = 0$. We have also kept the gauge dependence on ξ, previously omitted in Eq. (1.66).

Corrections to the gluon propagator are shown in Fig. 1.6 and are known as the vacuum polarization diagrams. Diagram 5 is zero in dimensional regularization; the combined contributions of the other diagrams are,

$$\begin{aligned}(\text{diagrams } 2+3+4) \;=\; & i\frac{\bar{g}^2}{16\pi^2}\delta_{ab}\left[\left(p^\mu p^\nu - g^{\mu\nu}p^2\right)\left(\frac{5}{3}C_A - \frac{4}{3}T_R N_f\right)\right. \\ & \left. + \frac{1}{2}C_A(1-\xi)\left(2p^\mu p^\nu - g^{\mu\nu}p^2\right)\right]\frac{1}{\varepsilon'} + \ldots . \qquad (1.70)\end{aligned}$$

The constants C_A and C_F are the group theory quantities defined in appendix A. The fermion propagating in the loop in diagram 2 could be any of potentially 6 quark flavours. We therefore sum this diagram over the number of active[7] fermion flavours, N_f. The Abelian, QED result can be obtained from Eq. (1.70) by setting all terms not involving N_f to zero and adopting the value of T_R appropriate to QED.

[7]We describe what we mean by 'active fermions' in section 5.3.



(diagram 6) (diagram 7)

Figure 1.7: One-loop corrections to the quark-gluon vertex.

Diagrams 6 and 7 are known as the vertex correction diagrams. Their combined regularized form is,

$$(\text{diagrams } 6+7) \;=\; ig\frac{\overline{g}^2}{16\pi^2}\gamma^\mu(T^a)_{ij}[4C_A+4C_F-(1-\xi)(C_A+4C_F)]\frac{1}{\varepsilon'}+\ldots. \tag{1.71}$$

Using dimensional regularization we have exposed the singularity structure of the one-loop correction diagrams of Figs. 1.5 - 1.7. Via the process of renormalization these singularities can now be systematically removed.

1.5.2 Renormalization constants

As was the case with ϕ^4 theory, we are required to redefine our Lagrangian in terms of bare and renormalized quantities, related to each other via renormalization constants. For $\mathcal{L}_{\text{QCD}}$ we perform the following redefinitions:

$$\begin{aligned} (G^\mu_a)_B &= Z_A^{1/2}(G^\mu_a)_R, & \psi_B &= Z_\psi^{1/2}\psi_R, \\ \eta^a_B &= Z_\eta^{1/2}\eta^a_R, & \xi_B^{-1} &= Z_\xi\xi_R^{-1}, \\ m_B &= Z_m m_R, & g_B &= Z_g g_R, \end{aligned} \tag{1.72}$$

where,

$$Z_i \;=\; 1+\Delta Z_i. \tag{1.73}$$

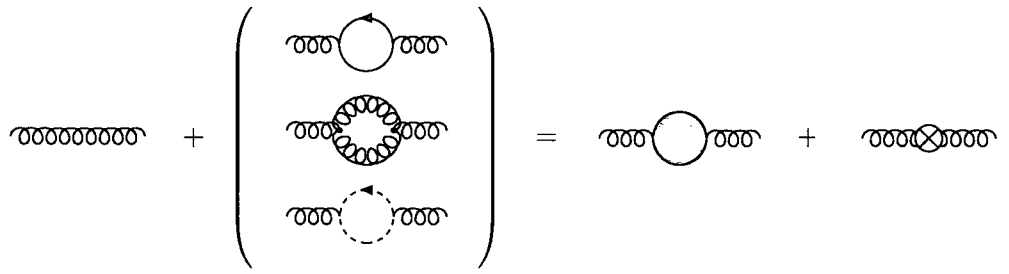


Figure 1.8: Comparison of 'bare' and renormalized diagrams for the gluon propagator, at one-loop. The first term on the RHS is the full propagator and contains no singularities. Hence, the singularity structure of ΔZ_i can be determined by comparing the LHS and RHS of this equation.

In massless QCD we can omit the equation for Z_m, leaving us with 5 renormalization constants. We redefine the Lagrangian of Eq. (1.15) as the bare Lagrangian and all the parameters within it as bare parameters, denoted by a subscript B. We also define the renormalized Lagrangian, $\mathcal{L}$, as in Eq. (1.15) but with all parameters being renormalized parameters (instead of writing all renormalized parameters with a subscript R we omit it and assume, from now on, that all parameters without a subscript are renormalized). The bare and renormalized Lagrangians are related to each other by,

$$\mathcal{L}_B = \mathcal{L} + \Delta\mathcal{L}, \tag{1.74}$$

where $\Delta\mathcal{L}$ is the *counterterm* Lagrangian. Substituting Eq. (1.72) into Eq. (1.74) and using Eq. (1.73) allows us to obtain $\Delta\mathcal{L}$ in terms of renormalized parameters and the renormalization constants,

$$\begin{aligned}\Delta\mathcal{L} &= \overline{\psi}_a\,(i\not{\partial} - m\Delta Z_m)_{ab}\,\psi_b\,\Delta Z_\psi - \frac{1}{4}\,F^a_{\mu\nu}F^{\mu\nu}_a\,\Delta Z_A \\ &\quad + \mathcal{L}_{\rm gf}\,(\Delta Z_A + \Delta Z_\xi) + \mathcal{L}_{\rm fp}\,\Delta Z_\eta \\ &\quad + \left(\Delta Z_g + \frac{\Delta Z_A}{2}\right)(\mathcal{L}_I^{q\text{-}fp} + \mathcal{L}_I^{q\text{-}g} + \mathcal{L}_I^{3g} + 2\mathcal{L}_I^{4g}).\end{aligned} \tag{1.75}$$

The renormalization constants encode the singularity structure of the loop diagrams. At one-loop they contain only simple poles, the residues of which are the crucial information obtained from the regularization procedure.

The renormalization constants can be determined by comparing Green's functions cal-

culated from the LHS and RHS of Eq. (1.74). For example, if we calculate the gluon propagator from $\mathcal{L}_{\rm B}$, we obtain the tree level propagator, plus the one-loop corrections of Eq. (1.70). $\mathcal{L}$ and $\Delta\mathcal{L}$ have their own set of Feynman rules. $\mathcal{L}$ leads to Green's functions written in terms of renormalized parameters, corresponding to what we actually measure. Also, $\Delta\mathcal{L}$ gives rise to a set of Feynman rules similar to those of Figs. 1.2 and 1.3, but which include extra factors in the form of combinations of ΔZ_i. The appropriate combinations of these are derived from Eq. (1.75). Diagrams constructed from these rules are known as counterterm diagrams and are represented by placing a $\otimes$ symbol on them. Comparing the singularity structure of Green's functions derived from each side of Eq. (1.74) allows us to obtain the poles in Z_i.

For the gluon propagator example, we can compare the one-loop corrections written in terms of bare parameters with the renormalized propagator and the counterterm propagator. This is shown in Fig. 1.8 and expressed mathematically below,

$$i\frac{\bar{g}^2\delta_{ab}}{24\pi^2}\frac{1}{\varepsilon}\left[(p^2g^{\mu\nu}-p^\mu p^\nu)\left(2T_RN_f-\frac{5}{2}C_A\right)-\frac{3}{4}C_A(1-\xi)(g^{\mu\nu}p^2-2p^\mu p^\nu)\right]$$
$$= i\Delta Z_A(-p^2g_{\mu\nu}+p_\mu p_\nu)\delta_{ab}-i(\Delta Z_A+\Delta Z_\xi)\xi^{-1}p_\mu p_\nu\delta_{ab}, \qquad (1.76)$$

where we have omitted the tree level gluon propagator from both sides. The appropriate combinations of ΔZ_i used in Eq. (1.76) can be obtained by determining which terms in Eq. (1.75) contribute to the gluon propagator. From Eq. (1.76), the renormalization constants Z_A and Z_ξ can be identified,

$$\Delta Z_A = -\frac{\bar{g}^2}{16\pi^2}\frac{1}{\varepsilon}\left[-\frac{5}{3}C_A+\frac{4}{3}T_RN_f-\frac{C_A}{2}(1-\xi)\right], \qquad (1.77)$$

$$\Delta Z_\xi = \frac{\bar{g}^2}{16\pi^2}\frac{1}{\varepsilon}\left[-\frac{5}{3}C_A+\frac{4}{3}T_RN_f-\frac{C_A}{2}(1-\xi^2)\right]. \qquad (1.78)$$

Repeating this procedure for the quark self-energy and then the ghost self-energy, one obtains,

$$\Delta Z_\psi = -\frac{\bar{g}^2C_F}{16\pi^2\varepsilon}\xi, \qquad \Delta Z_m = -\frac{\bar{g}^2C_F}{16\pi^2\varepsilon}(2+\xi), \qquad (1.79)$$

$$\Delta Z_\eta = -\frac{\bar{g}^2C_A}{16\pi^2\varepsilon}\left(\frac{3}{4}-\frac{5}{4}\xi\right). \qquad (1.80)$$

Finally, the one-loop quark-gluon vertex function can be used to obtain Z_g,

$$\Delta Z_g = -\frac{\bar{g}^2}{16\pi^2\varepsilon}\frac{1}{6}\left[11C_A - 4T_R N_f\right]. \tag{1.81}$$

Note that this expression is independent of ξ. Equations (1.77) - (1.81) represent a successful renormalization of QCD at one-loop and can be used to perform $\mathcal{O}(g^2)$ calculations of renormalized Green's functions.

To summarize, the singularities occurring at one-loop have been absorbed into the unphysical parameters of the bare Lagrangian. Our reward for this feat is a set of finite expressions for the one-loop quantum corrections to QCD Green's functions, and the ability to calculate perturbative corrections to observables, order-by-order, in the following form,

$$\mathcal{F}(\alpha_s) = \sum_{n=0}^{\infty} f_n \alpha_s^{n+1}, \tag{1.82}$$

where $a_s = g^2/4\pi$. However, this is at the cost of the introduction of an element of ambiguity, the origin of which is the arbitrariness of both the renormalization scale and of the finite subtractions made by Z_i. As we shall see in section 1.8, truncating Eq. (1.82) at fixed-order $\mathcal{O}\left(\alpha_s^n\right)$ leads to an $\mathcal{O}\left(\alpha_s^{n+1}\right)$ dependence of the result on the way in which it has been renormalized. Furthermore, when we attempt to resum leading logs in order to define the running coupling, this $\mathcal{O}\left(\alpha_s^{n+1}\right)$ dependence is absorbed into the lower-order terms in the truncated sum. Hence fixed-order predictions will depend on the way they are renormalized.

1.5.3 Renormalization scheme

As a result of the renormalization process, all singularities occurring at one-loop have been absorbed into the renormalization constants, and sense can finally be made of QCD Green's functions. This process can be generalized to higher-loop diagrams and consequently, QCD can be renormalized order-by-order in perturbation theory.

However, although the loop diagrams fix the pole structure of the renormalization constants, their finite part is arbitrary. We could shift the ΔZ_i's by any finite amount whilst still meeting the condition that renormalized Green's functions must be finite.

Furthermore, the renormalized one-loop diagrams all have a dependence on the arbitrary scale μ.

We are forced to choose what finite part of the loop diagrams we will keep when we perform the subtractions, and this choice is reflected in the value of the finite part of ΔZ_i. A particular set of choices for the finite parts of ΔZ_i, together with a choice of renormalization scale, defines the *renormalization scheme* (RS).

When we wrote down Eqs. (1.77) - (1.81), we set the finite parts of ΔZ_i to zero, and in so doing we unwittingly made a choice of renormalization scheme. We used the renormalization constants to subtract only the poles, and nothing else; this is known as the minimal subtraction (MS) scheme. We can transform from one scheme into another simply by performing a finite renormalization.

The two schemes used in this thesis are the modified minimal subtraction ($\overline{\text{MS}}$) scheme and the V-scheme. In $\overline{\text{MS}}$ the finite part $\gamma_{\text{E}} - \ln 4\pi$, present in Eqs. (1.68) - (1.71) and always present with DR, is removed. In the V-scheme the additional factor of $-5/3$, present in Eq. (1.67) is also removed. This scheme is useful when we come to consider chains of renormalized fermion loops, in chapter 3.

If we examine the effects of these finite renormalizations on the vacuum polarization function of Eq. (1.66), we see that they can simply be treated as different re-scalings of the renormalization scale. The renormalized form of Eq. (1.66) can be written,

$$-i\Pi(p^2) \;=\; i\frac{\overline{g}^2}{16\pi^2}\frac{4T_R}{3}\ln\left(\frac{-p^2}{\mu^2 \mathrm{e}^{C_{\text{RS}}}}\right), \tag{1.83}$$

where C_{RS} depends on the renormalization scheme being applied and in the three above mentioned cases, takes the values $C_{\text{MS}} = \ln 4\pi + 5/3 - \gamma_{\text{E}}$, $C_{\overline{\text{MS}}} = 5/3$ and $C_{\text{V}} = 0$. From this, it is easy to see how different choices of finite subtraction will affect the result of any physical observable calculated from the vacuum polarization function. This problem hinders attempts to unambiguously define QCD observables.

1.6 The renormalization group equation

It is clear that through the renormalization procedure, a dependence of physical observables on the renormalization scale μ is generated. This dependence can be defined more rigorously through the renormalization group equation. In order to show this we must first know how bare and renormalized OPI Green's functions relate to each other.

For QCD to be renormalizable, the divergences in the generating functionals must be removed; this has been achieved partially by the renormalization carried out via Eq. (1.72). However, this must be supplemented by renormalization of the source terms, which must satisfy,

$$J \cdot G = J_B \cdot G_B, \qquad \eta^\dagger \cdot \zeta = \eta_B^\dagger \cdot \zeta_B, \qquad \chi\overline{\psi} = \chi_B \overline{\psi}_B. \tag{1.84}$$

The bare and renormalized functionals are related by,

$$\mathcal{X}[J, \zeta, \zeta^\dagger, \chi, \overline{\chi}] = \mathcal{X}_B[J_B, \zeta_B, \zeta_B^\dagger, \chi_B, \overline{\chi}_B], \tag{1.85}$$

where $\mathcal{X}_B$ is simply the functional of Eq. (1.34) but written entirely in terms of bare parameters and $\mathcal{X}$ is the renormalized form of $\mathcal{X}_B$, obtained by using the Lagrangian $\mathcal{L}+\Delta\mathcal{L}$ in Eq. (1.53). It then follows (using Eqs. (1.35) and (1.36)) that renormalized OPI momentum space Green's functions are related to their bare counterparts by,

$$\tilde{G}^{(n)}(\mathbf{p}, \overline{g}, \xi, m, \mu) = Z_A^{n_G/2} Z_\psi^{n_\psi/2} \tilde{G}_B^{(n)}(\mathbf{p}, g_B, \xi_B, m_B). \tag{1.86}$$

Here, n_G and n_ψ are the number of external gluon and quark legs, respectively and $\mathbf{p}$ represents the set of external momenta, $\{p_1, p_2, \ldots, p_n\}$. The factor of $Z_A^{n_G/2} Z_\psi^{n_\psi/2}$ appears because for renormalized Green's functions, the functional differentiations in Eq. (1.54) are with respect to renormalized source terms.

Differentiating both sides of Eq. (1.86) with respect to μ, and then rearranging gives,

$$\left[\mu\frac{\partial}{\partial\mu} + \tilde{\beta}_{\overline{g}}\frac{\partial}{\partial\overline{g}} + \beta_\xi\frac{\partial}{\partial\xi} - \gamma_m m\frac{\partial}{\partial m} - n_G\gamma_G - n_\psi\gamma_\psi\right]\tilde{G}^{(n)} = 0. \tag{1.87}$$

This can be written more concisely as,

$$\left[\mu\frac{\partial}{\partial\mu} + \tilde{\beta}_{\overline{g}}\frac{\partial}{\partial\overline{g}} + \mathbb{D}\right]\tilde{G}^{(n)} = 0. \tag{1.88}$$

Here,

$$\mathbb{D} = \beta_\xi \frac{\partial}{\partial \xi} - \gamma_m m \frac{\partial}{\partial m} - n_G \gamma_G - n_\psi \gamma_\psi, \tag{1.89}$$

and the following quantities have been defined,

$$\tilde{\beta}_{\overline{g}} = \mu \frac{\partial \overline{g}}{\partial \mu,} \qquad \beta_\xi = \mu \frac{\partial \xi}{\partial \mu}, \qquad \gamma_m = -\frac{\partial \ln m}{\partial \ln \mu}, \tag{1.90}$$

$$\gamma_G = \frac{\mu}{2} \frac{\partial \ln Z_G}{\partial \ln \mu}, \qquad \gamma_\psi = \frac{\mu}{2} \frac{\partial \ln Z_\psi}{\partial \ln \mu}. \tag{1.91}$$

Equation (1.87) is known as the renormalization group equation (RGE). It demonstrates that changes in unrenormalized (bare) Green's functions, induced by varying the renormalization scale, are compensated for by changes in the other RS dependent parameters, $\overline{g}, \xi$ and m. $\tilde{\beta}_{\overline{g}}$ is known as the beta-function equation and is crucial in determining the energy dependence of QCD observables. It can be calculated as a perturbation series, in powers of $\bar{g}$, the leading order coefficient of which can be computed from the diagrams in Figs. 1.5 - 1.7. Using Eq. (1.65), we can write,

$$\begin{aligned} \tilde{\beta}_{\overline{g}} &= \frac{\partial}{\partial \ln \mu} \left(\frac{\mu^{-\epsilon} g_B}{Z_g} \right) \\ &= -\epsilon \overline{g} - \frac{2\Delta Z_g}{Z_g} \tilde{\beta}_{\overline{g}}, \end{aligned} \tag{1.92}$$

$$\Rightarrow \tilde{\beta}_{\overline{g}} = -\epsilon \overline{g} - \frac{b}{8\pi^2} \overline{g}^3 + \cdots, \tag{1.93}$$

where b is the first beta-function coefficient [25,26] which, using Eqs. (1.77) - (1.81), is found to be,

$$b = \frac{1}{6}(11 C_A - 2 N_f). \tag{1.94}$$

For the definition of C_A, see appendix A. In general, the beta-function equation is a perturbative expansion in powers of the coupling; for convenience, we redefine it as,[8]

$$\frac{\partial a}{\partial \ln \mu} = \beta(a) = -b a^2 (1 + ca + c_2 a^2 + c_3 a^3 + \cdots), \tag{1.95}$$

[8]There exist many different conventions for defining the beta-function equation, which differ in such ways as: the definition of a with respect to α, the use of g or α instead of a, the differential being with respect to μ^2 instead of μ etc. Consequently, our definitions of β, b, c etc. may differ from those in other works.

where,

$$a(\mu) = \frac{\alpha_s}{\pi} = \frac{g^2}{4\pi^2}. \tag{1.96}$$

The sub-leading coefficients can be obtained by calculating diagrams beyond one-loop and are found to be [27–31],

$$c = \frac{1}{12b}\left(-\frac{3}{2}C_A[7C_A + 11C_F] + 3b[5C_A + 3C_F]\right), \tag{1.97}$$

$$c_2^{\overline{\rm MS}} = \frac{2857 - \frac{5033}{9}N_f + \frac{325}{27}N_f^2}{64b}, \tag{1.98}$$

$$c_3^{\overline{\rm MS}} = \frac{\zeta_3 + \frac{149753}{6} - \left(\frac{6508}{27}\zeta_3 + \frac{1078361}{162}\right)N_f + \left(\frac{6472}{81}\zeta_3 + \frac{50065}{162}\right)N_f^2 + \frac{1093}{729}N_f^3}{256b}. \tag{1.99}$$

The coefficients b and c are RS invariant i.e. they are the same no matter which RS we use to calculate them. However, subsequent coefficients are all RS dependent, and hence Eqs. (1.98) and (1.99) reflect their values calculated in the $\overline{\rm MS}$ scheme. Equations (1.97) - (1.99) can also be rewritten as expansions in powers of b,

$$c = -\frac{107}{8\,b} + \frac{19}{4}, \tag{1.100}$$

$$c_2^{\overline{\rm MS}} = -\frac{37117}{768\,b} + \frac{243}{32} + \frac{325}{192}\,b, \tag{1.101}$$

$$c_3^{\overline{\rm MS}} = \frac{12185857 + 1389486\,\zeta_3}{13824\,b} - \frac{5857771 + 932400\,\zeta_3}{27648} + \frac{7761 + 1618\,\zeta_3}{576}\,b - \frac{1093}{6912}\,b^2. \tag{1.102}$$

1.7 The running coupling and asymptotic freedom

As a direct consequence of the RGE, the dependence of the coupling on μ is dictated by the beta-function equation,

$$\mu\frac{\partial a}{\partial \mu} = \beta(a). \tag{1.103}$$

Truncating Eq. (1.103) at leading order (LO) and integrating the resulting differential equation between the limits μ and Q gives,

$$b \ln \frac{Q}{\mu} = \frac{1}{a(Q)} - \frac{1}{a(\mu)}. \tag{1.104}$$

Rearranging this yields,

$$\frac{1}{a(Q)} - b \ln Q = \frac{1}{a(\mu)} - b \ln \mu \equiv -b \ln \Lambda. \tag{1.105}$$

The functions of μ and Q in the above equation are identical and therefore independent of both μ and Q. We identify Λ as the *universal scale parameter* of QCD; a Q and μ independent quantity which parameterizes the missing boundary condition in Eq. (1.103).

Rearranging Eq. (1.105) allows us to write a independently of μ:

$$a(Q) = \frac{a(\mu)}{1 + ba(\mu) \ln (Q/\mu)} \tag{1.106}$$

$$\equiv \frac{1}{b \ln (Q/\Lambda)}. \tag{1.107}$$

Λ is a physical QCD parameter, in the sense that by studying the strength of QCD interactions at different values of Q we could, in theory, obtain a value for Λ by fitting $a(Q)$ to the data. Indeed, it can be argued that Λ is more fundamental than even the coupling, because it can be defined in an RS-invariant way (see section 1.8.1).

It follows from Eqs. (1.68) - (1.71) that the measured (renormalized) coupling receives the following finite renormalizations due to the diagrams in Figs. 1.5 - 1.7,

$$a(Q) = a(\mu) - \frac{1}{2} b a^2(\mu) \left[\ln \left(\frac{Q^2}{\mu^2} \right) + \text{finite} \right] + \mathcal{O}\left(a^3(\mu) \ln^2 \left(\frac{Q^2}{\mu^2} \right) \right), \tag{1.108}$$

which can be written as a 'leading log' expansion in $a^{n+1} \ln^n$,

$$a(Q) = a(\mu) - \frac{1}{2} b a^2(\mu) \ln \left(\frac{Q^2}{\mu^2} \right) + \mathcal{O}\left(a^3(\mu) \ln^2 \left(\frac{Q^2}{\mu^2} \right) \right). \tag{1.109}$$

From this, we can see that expanding Eq. (1.106) as a Taylor series would reproduce the leading logarithmic components of Eq. (1.109). In effect, Eq. (1.107) has resummed the logarithms present in Eq. (1.109), and we say that the coupling of Eq. (1.107) is

'renormalization group (RG) improved'.

Via the beta-function equation, the coupling has gained a dependence on Q; because of this the coupling is said to be a 'running coupling'. From Eq. (1.107) we can see that the coupling, and therefore QCD observables, have the following important properties:

Asymptotic Freedom [25, 26] For $Q \gg \Lambda$, the coupling is small and has the limit $a(Q) \to 0$ as $Q \to \infty$. In this region, the perturbative approximation is on firm ground; we can safely use Eq. (1.82), and need only include the first few terms to make accurate predictions. Quarks with $Q \gg \Lambda$ will behave as if they are free particles.

Infrared slavery In the region $Q \approx \Lambda$ the coupling becomes very large and this corresponds to interactions becoming very strong. As a result of this, the convergence properties of Eq. (1.82) will suffer and the perturbative approximation becomes unreliable. The behaviour of the coupling at these low energies gives rise to the confinement of quarks inside baryons. Free quarks cannot exist as they can always lower their energy state by combining with virtual quarks created from the vacuum.

The value of Λ serves as a guide to the (energy) range of validity of perturbation theory. The existence of a singularity in the coupling at $Q = \Lambda$ (known as the Landau pole), would seem to spell disaster for QCD. However, the singularity is a pathology and signals the fact that perturbation theory breaks down and is invalid for values of $a \gtrsim 1$. As a consequence, we must be careful about the energy scales at which we apply perturbation theory.

Solutions to the beta-function equation

Equation (1.107) gives the energy dependence of the coupling at one-loop. To further increase the accuracy of our prediction we can include higher-order corrections to the beta-function equation. Integrating the full beta-function equation gives the following transcendental equation,

$$\ln \frac{\mu}{\Lambda} = \int_0^a \frac{dx}{\beta(x)} + \text{infinite constant}, \tag{1.110}$$

where we have assumed asymptotic freedom ($a(\mu) \to 0$ as $\mu \to \infty$). The only constraint on the infinite constant in Eq. (1.110) is that it must possess the same singularity structure as the integral. However, the particular choice of finite constant made corresponds to a particular choice for the definition of Λ. In this thesis, we make the following choice,

$$\text{infinite constant} \;=\; \int_0^\infty \frac{dx}{bx^2(1+cx)}. \tag{1.111}$$

This differs from the standard choice [32], but coincides with that used by Stevenson in Ref. [33]. As a consequence, our definition of Λ is not the standard one; the two definitions are related by,

$$\tilde{\Lambda} \;\equiv\; \Lambda_{\text{Stevenson}} \;=\; \left(\frac{2c}{b}\right)^{-\frac{c}{b}} \Lambda_{\text{conventional}}. \tag{1.112}$$

Truncating the beta-function equation after the nth term gives us the following transcendental equation,

$$b \ln \frac{\mu}{\tilde{\Lambda}} \;=\; \frac{1}{a^{(n)}} + c \ln \frac{ca^{(n)}}{1+ca^{(n)}} + \int_0^{a^{(n)}} dx \left[\frac{b}{\beta^{(n)}(x)} + \frac{1}{x^2(1+cx)}\right], \tag{1.113}$$

where $\beta^{(n)}(x)$ is the beta-function equation truncated at n terms, and $a^{(n)}$ is the nth order approximation of the coupling.[9]

We now describe several higher-order approximations to the running coupling. Truncating Eq. (1.103) at order a^3 and integrating the resultant differential equation gives us Eq. (1.113) with $n = 2$. In this case, the term in the integral vanishes and the solution to this equation can be written in closed form,

$$a_0(z(\mu)) \;=\; \frac{-1}{c\,[1 + W_{-1}(z(\mu))]}, \tag{1.114}$$

where,

$$z(x) \;=\; -\frac{1}{\mathrm{e}} \left(\frac{x}{\tilde{\Lambda}}\right)^{-b/c}, \tag{1.115}$$

and $W_{-1}(z)$ is the Lambert-W function, defined by the following transcendental equa-

[9] We can make the RS dependence of this coupling more explicit by writing it as $a^{(n)}(\mu, c_2, c_3, \ldots, c_n)$.

tion [34],

$$z = W(z)\mathrm{e}^{W(z)}. \tag{1.116}$$

The '-1' subscript denotes that we have chosen the -1 branch of $W(z)$; this is necessary in order to ensure asymptotic freedom [35,36]. We label the coupling in Eq. (1.114) with a '0' because, not only is it the second order approximation to the coupling, but it also corresponds to an RS in which all of the RS dependent coefficients of the beta-function equation are zero, i.e. $a_0 = a(\mu, 0, \ldots, 0)$; this scheme is known as the 't Hooft scheme. Because of this, we drop the $(n = 2)$ superscript because, although in any standard scheme it is a next-to-leading-order (NLO) approximation, in the 't Hooft scheme it is exact.

At NLO and next-to-next-to-leading-order (NNLO) ($n = 2$ and $n = 3$ in Eq. (1.113)) the coupling can be approximated by the following expansions in logs of μ/Λ:

$$a^{(2)}(\mu) = \frac{1}{\tau}\left(1 - \frac{c}{\tau}\log(2\tau/b)\right) + \mathcal{O}\left(\frac{1}{\tau^3}\right), \tag{1.117}$$

$$\begin{aligned} a^{(3)}(\mu) &= \frac{1}{\tau}\left(1 - \frac{c}{\tau}\log(2\tau/b)\right. \\ &+ \left.\left(\frac{c}{\tau}\right)^2\left[\log^2(2\tau/b) - \log(2\tau/b) + \frac{c_2}{c^2} - 1\right]\right) + \mathcal{O}\left(\frac{1}{\tau^4}\right). \end{aligned} \tag{1.118}$$

Here, $\tau \equiv b \ln \frac{\mu}{\Lambda}$, and these approximations are derived using the conventional definition of Λ in Eq. (1.112).

1.8 Renormalization scheme dependence

We noted in section 1.5.2 how renormalized perturbation series can depend on the choice of RS. It is clear that such dependence can disrupt the predictive power of QCD. Let us now put this analysis on a firmer footing.

Consider a generic, dimensionless QCD observable, $R(Q)$, dependent on a single dimensionful scale, Q. Perturbative corrections to $R(Q)$, denoted by a calligraphic letter,

can always be written,[10]

$$\mathcal{R}(Q) \;=\; a + \sum_{n=1}^{\infty} r_n a^{n+1}. \tag{1.119}$$

Physical observables *must* be RS invariant and therefore,

$$\frac{\partial}{\partial(\mathcal{RS})}\mathcal{R}(Q) \;=\; 0, \tag{1.120}$$

where we have used $\mathcal{RS}$ to symbolically represent variation of and dependence on renormalization scheme. Combining the two equations above gives,

$$\frac{\partial}{\partial(\mathcal{RS})}\left(\sum_{n=0}^{\infty} r_n(\mathcal{RS})a^{n+1}(\mathcal{RS})\right) \;=\; 0. \tag{1.121}$$

So although the coefficients and the coupling depend on the RS, this dependence cancels in the full expansion. In reality, we are not able to calculate the $n = \infty$ term in the series, and truncation of Eq. (1.121) at finite order leads directly to a dependence of the prediction on choice of RS. If we assume that we have an order $\mathcal{O}(a^N)$ calculation for $\mathcal{R}(a)$, i.e. we know the coefficients $r_{n<N}$, then we can relate the RS dependence of this result to the higher-order corrections. From Eq. (1.121),

$$\begin{aligned}\frac{\partial}{\partial(\mathcal{RS})}\left(\sum_{n=0}^{N-1} r_n(\mathcal{RS})a^{n+1}(\mathcal{RS})\right) \;&=\; -\frac{\partial}{\partial(\mathcal{RS})}\left(\sum_{n=N}^{\infty} r_n(\mathcal{RS})a^{n+1}(\mathcal{RS})\right) \\ &=\; \mathcal{O}(a^{N+1}).\end{aligned} \tag{1.122}$$

From this we can infer that the difference between $\mathcal{O}(a^N)$ calculations in two different schemes is of order $\mathcal{O}(a^{N+1})$. This is a direct consequence of demanding that perturbation theory be self-consistent [33].

So, further to the inaccuracy inherent in the truncation of an infinite series at finite order, we also have theoretical uncertainty, of the same order, generated by the non-uniqueness of the renormalization procedure. It is arguable as to which of these sources of uncertainty is more troubling. Indeed, as we shall see in later chapters, the known convergence properties of QCD perturbation series call into question the validity of a finite order truncation. However, the RS dependence problem is compounded by

[10] We can transform *any* perturbative result into this form by scaling, subtracting a constant, and raising to a suitable power.

the fact that the higher-order coefficients (which one might imagine could be used to estimate the theoretical uncertainty) are themselves RS dependent.

In an attempt to gain control of this source of ambiguity in QCD, we show in the next section how the RS dependence of perturbative expansions can be parameterized. We then proceed to describe three approaches to dealing with the RS dependence problem.

1.8.1 Parameterizing renormalization scheme dependence

When truncated at finite n, the perturbative expansion for $R(a)$ (given in Eq. (1.119)) possesses RS dependence from two sources: the coupling exhibits RS dependence due to its dependence on μ and c_i (expressed in the beta-function equation) and also, the perturbative coefficients r_n are RS dependent.

The beta-function equation can be rewritten (in the notation of Ref. [33]) as,

$$\frac{da}{d\tau} = -a^2(1 + ca + c_2a^2 + \cdots + c_ia^i + \cdots), \tag{1.123}$$

where $\tau \equiv b \ln \frac{\mu}{\Lambda}$. In order to study the effects of changes in RS, we consider the coupling defined in two different unspecified schemes, denoted by a and $\bar{a}$. Using Eq. (1.72) we can relate both a and $\bar{a}$ to the bare coupling,

$$a = Z_g^2 a_B, \qquad \bar{a} = \bar{Z}_g^2 a_B. \tag{1.124}$$

The pole structures of Z and $\bar{Z}$ must be identical and so a and $\bar{a}$ are related by only a finite renormalization, hence,

$$a(\mu) = \bar{a}(\mu)(1 + \nu_1\bar{a}(\mu) + \nu_2\bar{a}(\mu)^2 + \cdots). \tag{1.125}$$

Both a and $\bar{a}$ are defined at the same renormalization scale, μ, and the coefficients ν_i encode the difference between the different renormalization schemes. The two couplings can also be related to each other through their respective beta-function equations,

$$\overline{\beta}(\bar{a}) = \beta(a)\frac{d\bar{a}}{da}. \tag{1.126}$$

Consequently, the RS dependence of the coupling in these two different schemes can be related using their respective RS dependent coefficients, $\{c_2, c_3, \ldots c_n \ldots\}$ and

$\{\overline{c}_2, \overline{c}_3, \ldots \overline{c}_n \ldots\}$. In fact, as shown in Ref. [33], any particular RS can be characterized by the values of those coefficients in that scheme, together with μ/Λ (or equivalently τ). In other words, a particular RS corresponds to a particular set of values for $\{c_2, c_3, \ldots c_n \ldots\}$ and vice versa; we denote this set of parameters, $\{\mathcal{RS}\} \equiv \{\tau, c_2, c_3 \ldots c_n \cdots\}$. Because of this, $\{\mathcal{RS}\}$ can be used to parameterize all possible renormalization schemes, and hence the RS dependence of perturbative calculations.

The key to unlocking the RS dependence of the coupling lies in the beta-function equation. Differentiating Eq. (1.113) (with $n = \infty$) with respect to c_i yields,

$$\frac{\partial a}{\partial c_i} = -b\beta(a) \int_0^a \frac{x^{i+2}}{\beta^2(x)}\, dx. \tag{1.127}$$

We know that, given an $\mathcal{O}(a^n)$ calculation, the RS dependence of any observable is known to be zero up to order $n+1$, and hence we can write,

$$\frac{\partial \mathcal{R}^{(n)}}{\partial \tau} = \mathcal{O}\left(a^{n+1}\right), \qquad \frac{\partial \mathcal{R}^{(n)}}{\partial c_i} = \mathcal{O}\left(a^{n+1}\right). \tag{1.128}$$

Using the consistency of perturbation theory (expressed through Eq. (1.128)) in conjunction with the beta-function equation and Eq. (1.127), we can now obtain the explicit RS dependence of the coefficients r_n. We consider each r_n individually and assume in each case we have performed the relevant N^{n}LO calculation. For r_1 we have,

$$\mathcal{R}^{(1)} = a + r_1 a^2, \qquad \beta(a) = -ba^2(1 + ca), \tag{1.129}$$

$$\frac{\partial \mathcal{R}^{(1)}}{\partial \tau} = \mathcal{O}(a^2), \qquad \frac{\partial \mathcal{R}^{(1)}}{\partial c_i} = \mathcal{O}(a^2). \tag{1.130}$$

Performing the partial differentiations in Eq. (1.130) and equating coefficients gives,

$$\frac{\partial r_1}{\partial \tau} = 1, \qquad \text{and} \qquad \frac{\partial r_1}{\partial c_2} = 0, \tag{1.131}$$

and integrating these derivatives gives the explicit RS dependence of r_1,

$$r_1 = b \ln \frac{\mu}{\overline{\Lambda}} - X_0(Q). \tag{1.132}$$

Here $X_0(Q)$ is a constant of integration – by definition it is independent of all RS

parameters and hence is RS invariant. By contrast, the definition of μ and $\tilde{\Lambda}$ and the value of r_1 are RS-dependent. The fact that X_0 is RS invariant is remarkable and suggests that it is of genuine physical significance, in contrast to r_1, which is scheme dependent. Since $\mathcal{R}(Q)$ is a dimensionless function of a single scale Q, it follows that,

$$\begin{aligned} X_0(Q) &= \tau - r_1(\tau) \\ &\equiv b \ln \frac{Q}{\Lambda_{\mathcal{R}}}. \end{aligned} \tag{1.133}$$

Since X_0 and Q are RS invariant, using Eq. (1.133) we can infer that $\Lambda_{\mathcal{R}}$ is so too (although it is observable dependent). To obtain a relation between $\tilde{\Lambda}$ and $\Lambda_{\mathcal{R}}$ we consider Eq. (1.133) calculated in the $\overline{\text{MS}}$ scheme with $\mu = Q$. This leads to the following equation,

$$\Lambda_{\mathcal{R}} = \tilde{\Lambda}_{\overline{\text{MS}}} \exp\left(\frac{r_1^{\overline{\text{MS}}}(\mu = Q)}{b}\right). \tag{1.134}$$

Since $\Lambda_{\mathcal{R}}$ is RS invariant, specifying an RS does not affect the result. We can also use Eq. (1.134) to relate the scale parameters of two different schemes (RS and RS′) to each other:

$$\Lambda_{\text{RS}'} = \Lambda_{\text{RS}} \exp\left(\frac{r_1^{\text{RS}}(\mu) - r_1^{\text{RS}'}(\mu')}{b}\right). \tag{1.135}$$

The crucial point about Eq. (1.132) is that we can determine the explicit RS dependence of the perturbative coefficient r_1. Furthermore, we can extend this treatment to higher orders and determine the RS dependence of $r_{n>1}$. Performing the same procedure with r_2 yields,

$$\frac{\partial r_2}{\partial \tau} = 2r_1 + c, \qquad \frac{\partial r_2}{\partial c_2} = -1, \tag{1.136}$$

and for r_3,

$$\frac{\partial r_3}{\partial \tau} = 3r_2 + 2cr_1 + c_2, \qquad \frac{\partial r_3}{\partial c_2} = -2, \qquad \frac{\partial r_3}{\partial c_3} = -\frac{1}{2}. \tag{1.137}$$

Here we have used,

$$\frac{\partial a}{\partial c_2} = a^3 + \mathcal{O}(a^5), \qquad \frac{\partial a}{\partial c_2} = \frac{a^4}{2} + \mathcal{O}(a^5), \tag{1.138}$$

and $\frac{\partial r_n}{\partial c_i} = 0$ for $i > n$, which can be obtained from Eq. (1.127). The general form of the partial derivatives of r_n with respect to τ is,

$$\frac{\partial r_n}{\partial \tau} = \sum_{i=0}^{n} (n-i) c_i r_{n-i-1}. \tag{1.139}$$

By integrating up the partial derivatives in Eqs. (1.131), (1.136) and (1.137) we can obtain the dependence of r_n on the scheme parameters,

$$r_1 = \tau - X_0, \tag{1.140}$$

$$r_2 = r_1^2 + c r_1 - c_2 + X_2, \tag{1.141}$$

$$r_3 = 3 r_1 r_2 - 2 r_1^3 - \frac{c r_1^2}{2} + r_1 c_2 - \frac{1}{2} c_3 + X_3. \tag{1.142}$$

$$\vdots \quad \ddots$$

In the above equations, X_n are all generated as constants of integration with respect to the RS dependent parameters. Therefore they are by definition, RS-invariant quantities i.e. they have the same values no matter which scheme we use to calculate them. X_0 is of special significance due to its dependence on Q. The coefficient r_2 can be written as a polynomial in r_1,

$$r_2 = r_1^3 + \frac{5c}{2} r_1^2 + (3X_2 - 2c_2) r_1 - \frac{c_3}{2} + X_3, \tag{1.143}$$

and r_n has the general form,

$$r_n = \hat{r}_n + X_n - \frac{c_n}{n-1}. \tag{1.144}$$

Here, $\hat{r}_n$ is simply a polynomial in r_1, and is RG predictable from only an $\mathrm{N}^{n-1}\mathrm{LO}$ calculation of $\mathcal{R}(a)$. The coefficients X_n and c_n are obtainable by performing $\mathrm{N}^n\mathrm{LO}$ calculations of $\mathcal{R}(a)$ and $\beta(a)$ respectively.

We have demonstrated how to quantify the RS dependence of a general $\mathrm{N}^n\mathrm{LO}$ calculation. We have also seen how to make the dependence on the parameters $\{\mathcal{RS}\}$ explicit and how in doing this, RS invariant quantities appear naturally. We now proceed to discuss three different approaches to solving the RS dependence problem.

1.8.2 The physical scale approach

The physical scale (PS) approach derives its name from the fact that it involves setting the μ to some scale with physical meaning, such as the c.m. energy, Q. This is supplemented by the choice of some specific RS such as $\overline{\text{MS}}$. Such an approach falls into the category of approaches which propose that there is some 'best choice' of RS.

The motivation for the choice of scale $\mu = Q$ is the fact that at large orders of perturbation theory, the coefficients have the form,

$$r_n \sim \sum_{i=0}^{n} K_{ni} \left(b \ln \frac{\mu}{Q} \right)^i . \tag{1.145}$$

Thus, setting $\mu = Q$ allows us to control the (potentially large) logarithms present. The act of setting $\mu = Q$ seems, at first, dubious and unphysical. However, since μ is entirely arbitrary, it can be set to any value we wish; the error associated with different choices of μ is simply hidden inside the ambiguity inherent in RS dependent calculations. This error can be estimated by varying μ within the range $Q/2 < \mu < 2Q$.

The estimate of the higher-order coefficients in Eq. (1.145) is based on a simple LO analysis and assumes that RS dependence is given entirely by μ. A more sophisticated NLO analysis reveals that they have the following, more complicated form,

$$r_n \sim \sum_{i=0}^{n} \bar{K}_{ni} \left(b \ln \frac{\mu}{Q} \mathrm{e}^{r_1^{\mathrm{RS}}/b} \right)^i , \tag{1.146}$$

where $r_1^{\mathrm{RS}} = r_1^{\mathrm{RS}}(\mu = Q)$. The coefficients $\bar{K}_{ik}$ are independent of μ whereas K_{ik} are μ-dependent [37]. Hence, from Eq. (1.146) we can deduce that a more natural choice of renormalization scale is,

$$\mu = Q \, \mathrm{e}^{-r_1^{\mathrm{RS}}/b}. \tag{1.147}$$

In subsections 1.8.3 and 1.8.4 we will give a slightly different motivation for this choice of renormalization scale. However, the motivation for choosing the $\overline{\text{MS}}$ scheme with a physical scale is merely that it is convenient and widely used. Beyond that it is difficult to argue why it would be the 'best' scheme to use. Clearly, the ideal scheme is that in which the higher-order corrections (which quantify the RS dependence of fixed-order calculations) are maximally suppressed. However, the higher order terms

being unknown makes it difficult to assess the success of any particular scheme on such a basis.

1.8.3 Complete renormalization group improved perturbation theory

Through Eq. (1.95) the coupling, and therefore all QCD observables, gain a dependence on logarithms of $\mu/\tilde{\Lambda}$, with $\tilde{\Lambda}$ being RS dependent. This is in direct contradiction with the fact that all observables must be RS independent. In reality, the energy dependence of the coupling/observables must be built from physical logarithms, $\log Q/\Lambda_{\mathcal{R}}$, with $\Lambda_{\mathcal{R}}$ being RS invariant [37].

The dependence on unphysical logs (of $\mu/\tilde{\Lambda}$) is present because the standard RG improvement (described in section 1.7) is incomplete. It omits a set of RG predictable logarithms which could, in theory, be resummed; consequently, there remains a residual dependence on these unphysical logs. In Refs. [38,39] it was shown that by resumming *all* RG predictable physical logs, one can complete the RG improvement and build the $\log Q/\Lambda_{\mathcal{R}}$ dependence automatically, without having to impose a relation between μ and Q. Importantly, this procedure is independent of μ.

We begin by writing the perturbation series for $\mathcal{R}(a)$ in terms of the set of RS parameters $\{\mathcal{RS}\} = \{r_1, c_2, c_3, \ldots\}$. We have swapped the parameter τ for r_1, which is permitted since they both parameterize the LO RS dependence equally well (see Eq. (1.133)). Substituting the expressions for r_n given in Eqs. (1.140) - (1.142) into Eq. (1.119) gives,

$$\begin{aligned} \mathcal{R}(Q) &= a + r_1 a^2 + (r_1^2 + c r_1 + X_2 - c_2)a^3 \\ &+ (r_1^3 + \frac{5}{2} c r_1^2 + (3X_2 - 2c_2)r_1 + X_3 - \frac{1}{2}c_3)a^4 + \ldots . \end{aligned} \tag{1.148}$$

We then adopt the principle that we should resum all known RG predictable parts of Eq. (1.148) (i.e. the $\hat{r}_n$ terms in Eq. (1.144)) and we define this subset of terms as a_0. If we have an NLO calculation, then we know r_1 but not $X_2, X_3 \ldots$, and a_0 therefore has the form,

$$a_0 = a + r_1 a^2 + (r_1^2 + c r_1 - c_2)a^3 + (r_1^3 + \frac{5}{2} c r_1^2 - 2c_2 r_1 - \frac{1}{2}c_3)a^4 + \ldots . \tag{1.149}$$

This sum includes all terms in the full perturbative expansion which are RG-predictable at NLO i.e. those that don't include X_2, X_3 etc. Importantly, we can infer that a_0 is RS invariant, via the following argument. The expression in Eq. (1.148), when summed to all orders, is RS invariant. This sum differs from a_0 only by an RS invariant set of terms built from X_n, and therefore a_0 must also be RS invariant.

We have learnt that the values of the coefficients c_n, together with r_1, can be used to define any particular RS. Also, given that a_0 is RS invariant, calculating it in any scheme will yield the same result. We can take advantage of these facts by defining a special scheme in which all sub-leading coefficients of the beta-function equation as well as r_1 are set to zero, i.e. $r_1 = c_2 = c_3 = \ldots = c_n \ldots = 0$. This is known as the 't Hooft scheme [40] cf. Eq. (1.114). In a sense, this is moving in the opposite direction to the usual way in which an RS is defined. Previously, we made a set of finite subtractions using ΔZ_i and observed the effects of this on the values of the coefficients c_n and r_1. Now we are fixing the values of these coefficients first; the exact nature of the finite subtractions in ΔZ_i can then be derived from that information.

In the 't Hooft scheme, Eq. (1.149) reduces to $a_0 = a$, and therefore a_0 is simply the coupling calculated in that scheme. Hence, we can use Eq. (1.113), and then Eq. (1.133) to obtain,

$$\frac{1}{a_0} + c \ln \frac{c a_0}{1 + c a_0} = b \ln \frac{\mu}{\tilde{\Lambda}} \tag{1.150}$$

$$= b \ln \frac{Q}{\Lambda_{\mathcal{R}}}. \tag{1.151}$$

And using the results obtained in section 1.7, we can write the solution to this equation in closed form,

$$a_0(Q) = \frac{-1}{c[1 + W(z(Q))]}, \tag{1.152}$$

with

$$z(Q) = -\frac{1}{\mathrm{e}} \left(\frac{Q}{\Lambda_{\mathcal{R}}} \right)^{-b/c}. \tag{1.153}$$

We conclude that the observable at NLO has the following *complete renormaliza-*

tion group improved (CORGI) form,

$$\mathcal{R}(Q) \;=\; a_0(Q). \tag{1.154}$$

This treatment can easily be extended to higher-orders. At NNLO, we have the coefficient X_2 at our disposal. Hence we can resum a further set of RG-predictable terms in Eq. (1.148) which have the form,

$$a_0^3 X_2 \;=\; X_2 a^3 + 3X_2 r_1 a^4 + \ldots, \tag{1.155}$$

and consequently, the observable has the NNLO form,

$$\mathcal{R}(Q) \;=\; a_0(Q) + X_2 a_0^3. \tag{1.156}$$

Extending beyond NNLO, we finally obtain,

$$\tilde{\mathcal{R}}(Q) \;=\; a_0(Q) + \sum_{n=2}^{\infty} X_n a_0^{n+1}. \tag{1.157}$$

From now on, we label observables expressed in the CORGI representation with a tilde, in order to differentiate them from the standard perturbative representation. The coefficients X_n can be obtained from Eqs. (1.141) and (1.142). Since they are scheme invariant, we can use any scheme we like to evaluate them. For example,

$$X_2 \;=\; r_2 - r_1^2 - c r_1 + c_2 \Big|_{\overline{\rm MS},\,\mu=Q} \tag{1.158}$$

$$X_3 \;=\; r_3 - 3r_1 r_2 + 2r_1^3 + \frac{c r_1^2}{2} - r_1 c_2 + \frac{1}{2} c_3 \Big|_{\overline{\rm MS},\,\mu=Q} \tag{1.159}$$

where all coefficients on the RHS are calculated in $\overline{\rm MS}$ with $\mu = Q$.

It is curious that although Eq. (1.154) is an NLO result, it contains no a_0^2 term. Indeed, this term is also absent from the NNLO and all-orders results. However, closer inspection reveals that the a_0^2 term *is* present in Eq. (1.154), via the definition of $\Lambda_{\mathcal{R}}$. If we temporarily relabel a_0 with the type of scale used to define it,

$$a_0(Q,\Lambda) \;=\; \frac{-1}{c\,[1 + W(z(Q,\Lambda))]}, \qquad \text{where,} \qquad z(Q,\Lambda) \;=\; -\frac{1}{\mathrm{e}} \left(\frac{Q}{\Lambda}\right)^{-b/c}. \tag{1.160}$$

then $a_0(Q, \Lambda_{\mathcal{R}})$ is equal to the following expansion,

$$a_0(Q, \Lambda_{\mathcal{R}}) = a_0(\mu, \Lambda) + r_1^{\overline{\rm MS}} a_0(\mu, \Lambda)^2 + \mathcal{O}(a^3). \tag{1.161}$$

The above equation can be verified by substitution into Eq. (1.150). Thus, we see that the NLO information in Eq. (1.157) is encoded in $\Lambda_{\mathcal{R}}$. Rather than including this information in the form of an a_0^2 term, the CORGI expressions use $\Lambda_{\mathcal{R}}$ to resum *it* together with a further infinite subset of RG predictable terms.

Although $\Lambda_{\mathcal{R}}$ is in itself a valid parameter of the theory, it is often more convenient to relate it to an observable independent and more commonly recognized parameter, such as $\Lambda_{\overline{\rm MS}}$. We can convert between these two scales using Eqs. (1.134) and (1.112),

$$\Lambda_{\overline{\rm MS}} = \left(\frac{2c}{b}\right)^{c/b} \exp\left(-\frac{r_1^{\overline{\rm MS}}(\mu = Q)}{b}\right) \Lambda_{\mathcal{R}}. \tag{1.162}$$

We have obtained a result which is written purely in terms of RS invariant components and is therefore independent of the scheme we use in order to calculate it. This is a significant development, given the shaky ground we were on when we had to opt for some 'best' scheme. However, it must be noted that Eq. (1.157) *does* depend on the set of parameters we chose to parameterize the RS invariance, i.e. $\{r_1, c_2, c_3, \ldots\}$. Though the choice we have made is obviously natural, it is not unique, and there is an ambiguity in Eq. (1.157) associated with this freedom of choice. In a sense, we have traded RS dependence – with no obvious choice for the correct RS – for 'parameterization dependence', but with an obvious natural choice for the correct 'parameterization scheme'.

Indeed, it is argued in Refs. [38,39] that r_1 has special status, since it translates between μ and Q dependence, via Eq. (1.140). Using this equation, we can write,

$$r_1(\tau) = b \ln \frac{\mu}{\tilde{\Lambda}} - b \ln \frac{Q}{\Lambda_{\mathcal{R}}}. \tag{1.163}$$

To simplify the following demonstration, we set $c = 0$ and choose the 't Hooft scheme where $c_2 = c_3 = \ldots = 0$. We can write the coupling as,

$$a(\mu) = \frac{1}{b \ln(\mu/\tilde{\Lambda})}, \tag{1.164}$$

and at NLO, the RG improvement of Eq. (1.148) becomes,

$$\mathcal{R}(Q) \approx a(\mu) + r_1 a(\mu)^2 + r_1^2 a(\mu)^3 + \ldots. \tag{1.165}$$

This can be resummed, and then simplified using Eq. (1.163),

$$\begin{aligned} \mathcal{R}(Q) &\approx \frac{a(\mu)}{1 - r_1 a(\mu)} \\ &= \frac{a(\mu)}{1 - \left(b \ln \frac{\mu}{\tilde{\Lambda}} - b \ln \frac{Q}{\Lambda_{\mathcal{R}}}\right) a(\mu)} \\ &= \frac{1}{b \ln(Q/\Lambda_{\mathcal{R}})}. \end{aligned} \tag{1.166}$$

Thus we can see explicitly how the unphysical logs are 'eaten' by r_1 and the physical $\log(Q/\Lambda_{\mathcal{R}})$ dependence is built automatically.

Via the CORGI approach we have resummed a set of RG predictable $\log \mu/\Lambda$ terms, omitted by the standard PS prediction. Because of this, the uncertainty due to the freedom in choice of μ present in the PS result, is absent in the CORGI equivalent. Hence applying CORGI perturbation theory to phenomenological analyses will yield results for which the uncertainty due to RS dependence is heavily suppressed.

1.8.4 Effective charge scheme

We have learnt that the values of the coefficients c_n, together with r_1, can define any particular RS. The effective charge (EC) scheme [41,42] takes advantage of this fact by defining a scheme in which all sub-leading coefficients of the perturbative expansion of $\mathcal{R}$ are set to zero, i.e. $r_1 = r_2 = \ldots = r_n \ldots = 0$. As a result, the observable is simply equal to the coupling (defined in this scheme),

$$\mathcal{R}(Q) = a^{\mathrm{EC}}. \tag{1.167}$$

As a direct consequence of setting $r_1 = 0$ we can determine a relationship between μ and Q. Using Eqs. (1.132), (1.133) and then Eq. (1.134), we obtain,

$$\mu = \frac{\tilde{\Lambda}}{\Lambda_{\mathcal{R}}} Q$$

$$= Q\exp\left(-\frac{r_1^{\overline{\rm MS}}(\mu=Q)}{b}\right). \tag{1.168}$$

The energy dependence of $\mathcal{R}(a)$ is now directly related to the beta-function equation, and using Eq. (1.168) we can write,

$$\frac{\partial\mathcal{R}(Q)}{\partial\ln Q} = \beta^{\rm EC}(\mathcal{R}(Q)) \equiv \rho(\mathcal{R}(Q)), \tag{1.169}$$

where $\rho(a)$ is the beta-function equation in the EC scheme; because $\rho(a)$ is a beta-function its first two coefficients are known,

$$\rho(a) = -ba^2(1+ca+\rho_2a^2+\rho_3a^3+\ldots). \tag{1.170}$$

In order to obtain the higher-order coefficients we need to compare the EC scheme with some other scheme. Using Eq. (1.126) with the EC scheme as the barred scheme and the non-barred scheme being 'any other scheme', we obtain,

$$\rho(\mathcal{R}) = \beta(a(\mathcal{R}))\frac{d\mathcal{R}}{da}, \tag{1.171}$$

where $a(\mathcal{R})$ is an inversion of the perturbation series in Eq. (1.119). In the above non-barred scheme, it is given by,

$$a(\mathcal{R}) = \mathcal{R} - r_1\mathcal{R}^2 + \left(2r_1^2 - r_2\right)\mathcal{R}^3 + \cdots. \tag{1.172}$$

Expanding Eq. (1.171) in powers of $\mathcal{R}$ (using Eqs. (1.172) and (1.169)), and then equating coefficients, allows us to obtain ρ_n in terms of scheme dependent parameters,

$$\rho_2 = c_2 + r_2 - cr_1 - r_1^2, \tag{1.173}$$

$$\rho_3 = c_3 + 2r_3 - 4r_1r_2 - 2r_1\rho_2 - cr_1^2 - cr_1^2 + 2r_1^3. \tag{1.174}$$

$$\vdots \quad \ddots$$

By comparing this with Eqs. (1.140) - (1.142) we can see that the coefficients ρ_n are simply the RS invariants X_n, up to a factor of $1/(n-1)$ for $n = 2$ and 3. The relationship between X_4 and ρ_4 is slightly more complicated.

The Q dependence of the effective charge perturbative prediction now lies in a single power of the coupling viz Eq. (1.167). Solving the beta-function equation in the effective

charge scheme gives,

$$b \ln \frac{Q}{\Lambda_{\mathcal{R}}} = \int_{\mathcal{R}(Q)}^{\infty} \frac{dx}{x^2(1+cx)} + \int_0^{\mathcal{R}(Q)} dx \left[\frac{b}{\rho(x)} + \frac{1}{x^2(1+cx)} \right]. \quad (1.175)$$

This can be used as a means by which we may extract a value for Λ from a set of experimental data for $\mathcal{R}(Q)$. Alternatively, Eq. (1.175) can be used to make predictions for $\mathcal{R}(Q)$; exponentiating it yields,

$$\frac{\Lambda_{\mathcal{R}}}{Q} = \mathcal{F}(\mathcal{R}(Q))\,\mathcal{G}(\mathcal{R}(Q)). \quad (1.176)$$

Here, we have defined,

$$\mathcal{F}(\mathcal{R}) = \mathrm{e}^{-1/b\mathcal{R}} \left(1 + \frac{1}{c\mathcal{R}}\right)^{c/b}, \quad (1.177)$$

$$\mathcal{G}(\mathcal{R}) = \exp\left[-\int_0^{\mathcal{R}(Q)} dx \left(\frac{1}{\rho(x)} + \frac{1}{bx^2(1+cx)} \right) \right]. \quad (1.178)$$

We can then solve Eq. (1.176) iteratively to obtain $\mathcal{R}(Q)$ as a function of Q. $\Lambda_{\mathcal{R}}$ can then be obtained by obtaining the best fit to the data. Using Eq. (1.168), we can then extract $\Lambda_{\overline{\mathrm{MS}}}$ directly bin-by-bin from data for $\mathcal{R}(Q)$.

1.9 Non-perturbative corrections

We include here a short section on non-perturbative contributions to QCD observables. So far we have assumed that an observable can be reconstructed from its Taylor series, about the point $a = 0$. However, this assumption turns out to be false, as we cannot rule out the possibility of $\mathcal{R}(Q)$ containing terms of the form,

$$\mathcal{R}(Q) \ni \mathrm{e}^{-K/a(Q)}. \quad (1.179)$$

It is impossible to form a Taylor series from such terms; in this sense they are invisible to perturbation theory. Terms of this kind are known as non-perturbative corrections and in general take the form of inverse powers of the energy scale Q^2 (this can be seen by substituting the one-loop form of the coupling into Eq. (1.179)). Consequently, a

QCD observable will receive contributions from power corrections of the following form,

$$\mathcal{R}_{\rm NP}(Q) \;=\; \sum_{i=1}^{\infty} C_i \left(\frac{\Lambda^2}{Q^2}\right)^i . \tag{1.180}$$

Non-perturbative corrections are somewhat harder to evaluate than their perturbative counterparts. The operator product expansion (OPE) [43–46] is a commonly used approach and in the following section we briefly describe its origin and application.

1.9.1 The Wilson operator product expansion

In quantum field theory, one is often interested in evaluating the time ordered product of two operators,

$$T[A(x)B(y)]. \tag{1.181}$$

Often this quantity is not well-defined in the limit $x \to y$. However, the singular nature of this limit can be quantified by expanding Eq. (1.181) in a series of operators, $O_i(x)$, constructed from the components of the Lagrangian,

$$T[A(x)B(y)] \;=\; \sum_i C_i(x-y)\, O_i\left(\frac{x+y}{2}\right). \tag{1.182}$$

The operators $O_i(x)$ are the set of all non-singular local renormalized composite operators of the theory, with quantum numbers matching those of the product $A(x)B(y)$. The sequence of terms in the sum in Eq. (1.182) is organized in order of the canonical dimension of the operators. These composite operators are regular and hence the singularity structure of $T[A(x)B(y)]$ is wholly contained in the *Wilson coefficients*, C_i. Furthermore, in a theory with no dimensionful parameters, each C_i involves only a single power of $x - y$, obtainable by applying dimensional analysis to Eq. (1.182). In QCD they will be subject to corrections in the form of logarithms of the 'renormalization scale'. We use inverted commas because it is not meant in the same sense as the renormalization scale defined in section 1.5. We give more details of this in chapter 2.

Equation (1.182) is dominated by those terms which are most singular in x. Furthermore, there will only be a finite number of Wilson coefficients that are singular at $x = y$ and hence we can approximate Eq. (1.181) by considering only the first term (or first

few terms) in Eq. (1.182). This makes the OPE a powerful tool with which to study QFT.

The OPE technique has been applied to the calculation of both perturbative and non-perturbative corrections. Non-perturbative corrections can be evaluated by organizing the Wilson coefficients into a series in powers of $1/Q^2$ and then evaluating the relevant operators. The OPE also has an important perturbative application in the study of deep inelastic scattering and we study this in more detail in chapter 2.

1.10 QCD observables of interest

We now proceed to define several observables used in this thesis. We always work in massless QCD . This means that at a particular energy scale Q, we ignore the existence of quarks with masses $m_q > Q$ and treat all quarks with $m_q < Q$ as massless. This is implemented in perturbative expansions by treating the value N_f appearing in the coefficients, as a (step-like) function of energy e.g. it takes the values of 6 for $Q \geq m_t$, 5 for $m_t > Q \geq m_b$, etc. We will give more details of how this is implemented in chapter 5.

1.10.1 The Adler D-function

The Adler D-function is related to the vacuum polarization function, formally defined as the transverse part of the correlator of two (electromagnetic) vector currents $j_\mu(x)$, (with $i = u, d, s, \ldots$) in the Euclidean region ($Q^2 = -q^2$),

$$(q_\mu q_\nu - g_{\mu\nu}q^2)\Pi(Q^2) \;=\; 16\pi^2 i \int d^4x\, e^{iq.x}\langle 0|T[j_\mu(x)j_\nu(0)]|0\rangle. \qquad (1.183)$$

The currents, $j_\mu(x)$ have the form,

$$j_\mu(x) \;=\; \sum_i Q_i \overline{\psi}_i \gamma_\mu \psi_i. \qquad (1.184)$$

There exist several different definitions of $\Pi(Q^2)$ and the Adler D-function in the literature, all differing by constant factors. We define the Adler D-function as follows:

$$D(-Q^2) \;=\; -\frac{3}{4}Q^2\frac{d}{dQ^2}\Pi(Q^2). \qquad (1.185)$$

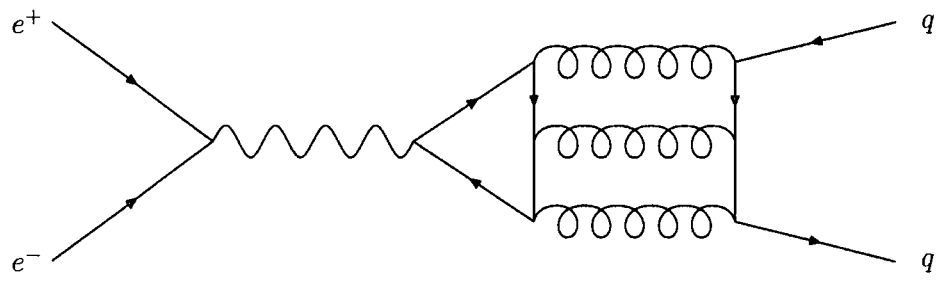


Figure 1.9: An example of a light-by-light scattering diagram. The two gluon version of the above diagram cancels and hence this is the lowest order light-by-light contribution.

Because $D(Q^2)$ is defined for space-like momenta only, it is known as a Euclidean observable. Unlike $\Pi(Q^2)$, the Adler D-function is an RG-invariant quantity. Hence we can consider $D(Q^2)$ to be a physical 'observable', despite the fact that it cannot be measured directly in experiment. Experimental values for $D(Q^2)$ can only be obtained indirectly, e.g. via the $R_{e^+e^-}$ ratio [47,48].

$D(Q^2)$ can be split into a constant term, known as the parton model result (also known as the Born term), and perturbative corrections,[11]

$$D(Q^2) = N_c \sum_f Q_f^2 \left(1 + \frac{3}{4} C_F \mathcal{D}(a)\right) + \left(\sum_f Q_f^2\right)^2 \tilde{\mathcal{D}}(a). \tag{1.186}$$

The $\tilde{\mathcal{D}}(a)$ term arises from so-called 'light-by-light' scattering diagrams, an example of which is given in Fig. 1.9. These contributions are suppressed relative to $\mathcal{D}(a)$ and therefore we neglect them. The remaining perturbative corrections take the form,

$$\mathcal{D}(a) = a + \sum_{n=1}^{\infty} d_n a^{n+1}. \tag{1.187}$$

The coefficients d_n are RS-dependent; d_1 and d_2 have been calculated in $\overline{\text{MS}}$ with $\mu = Q$, in [49–51] and [52,53] respectively. Their values are found to be,

$$d_1 = \left(-\frac{11}{12} + \frac{2}{3}\zeta_3\right) N_f + C_A \left(\frac{41}{8} - \frac{11}{3}\zeta_3\right) - \frac{1}{8} C_F, \tag{1.188}$$

[11] In this thesis, wherever we use a capital letter to denote an observable, we use the calligraphic version of that letter to represent the perturbative corrections.

$$
\begin{aligned}
d_2 &= \left(\frac{151}{162} - \frac{19}{27}\zeta_3\right) N_f^2 + C_A \left(-\frac{970}{81} + \frac{224}{27}\zeta_3 + \frac{5}{9}\zeta_5\right) N_f \\
&+ C_F \left(-\frac{29}{96} + \frac{19}{6}\zeta_3 - \frac{10}{3}\zeta_5\right) N_f + C_A^2 \left(\frac{90445}{2592} - \frac{2737}{108}\zeta_3 - \frac{55}{18}\zeta_5\right) \\
&+ C_A C_F \left(-\frac{127}{48} - \frac{143}{12}\zeta_3 + \frac{55}{3}\zeta_5\right) + C_F^2 \left(-\frac{23}{32}\right), \qquad (1.189)
\end{aligned}
$$

where ζ_n is the Riemann zeta function, defined in section B.4.

In addition to the perturbative corrections in Eq. (1.186), the Adler D-function will also have non-perturbative corrections and these can be studied within the OPE. Applying Eq. (1.182) to $D(Q)$, via Eq. (1.183), gives the following representation,

$$
D(Q^2) = \sum_n \sum_{\dim O_i = 2n} \frac{C_i(Q^2, \mu^2)\langle O_i(\mu^2)\rangle}{Q^{2n}}. \qquad (1.190)
$$

The relevant operators in this case are local gauge invariant operators constructed from the field operators of $\mathcal{L}_{\rm QCD}$. The $n = 0$ term simply corresponds to the perturbative corrections $\mathcal{D}(a)$. For $n = 1$ the relevant operator is constructed from the running quark masses $m_i(\mu)m_j(\mu)$. These will generate corrections of the form $m_i(\mu)m_j(\mu)/Q^2$, and will build the quark mass dependence that we have ignored by working in massless QCD. Non-perturbative corrections are generated by terms for which $n \geq 2$. At $n = 2$ the relevant operators are $\langle m_j\overline{\psi}_i\psi_i\rangle$ and $\langle FF\rangle$, known as the quark and gluon condensates respectively. These non-perturbative corrections are evaluated by separating long and short distance effects at a *factorization scale* μ. The long-distance effects can be absorbed into a non-perturbative parameter which can be determined by fitting to experimental data and the short distance effects are then calculated perturbatively as an expansion in $a(Q)$. Correspondingly, $C_i(Q^2, \mu^2)$ can be expressed as a function of the coupling.

1.10.2 The $R_{e^+e^-}$ ratio

Another good test of QCD is found in the production of hadrons from e^+e^- annihilation. This process occurs via the production of a quark anti-quark pair and is depicted in Fig. 1.10. The relationship between this process and QCD is simpler if we express it

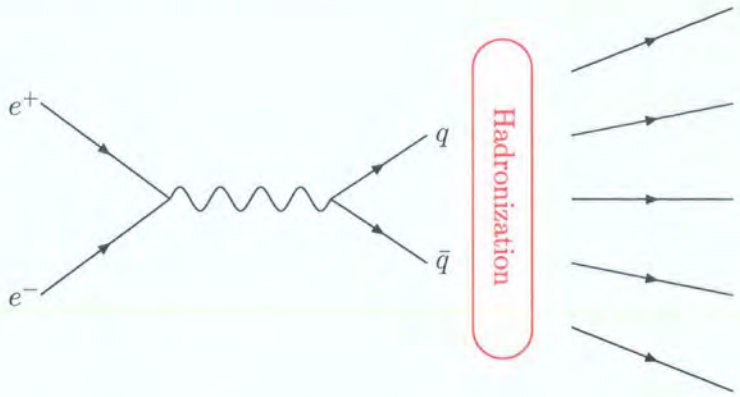


Figure 1.10: Depiction of the process $e^+e^- \to$ hadrons.

via the following ratio of cross sections,

$$R_{e^+e^-} \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}. \tag{1.191}$$

The hadronization process depicted in Fig. 1.10 is strictly a non-perturbative effect and this fact would appear to complicate the evaluation of the cross section. However, the details of hadronization are irrelevant since the q-$\bar{q}$ pair are converted into hadrons with a probability of 1. Consequently, we can make the identification,

$$R_{e^+e^-} = \frac{\sum_q \sigma(e^+e^- \to q\bar{q})}{\sigma(e^+e^- \to \mu^+\mu^-)}, \tag{1.192}$$

and the LO result for this can be obtained simply by counting the number of possible $q\bar{q}$ states, weighted by their charges,

$$R_{e^+e^-}(s) = N_c \sum_q Q_q^2. \tag{1.193}$$

This result is subject to perturbative corrections of the form,

$$R_{e^+e^-}(s) = N_c \sum_q Q_q^2 \left(1 + \frac{3}{4} C_F \mathcal{R}(a)\right) + \left(\sum_q Q_q\right)^2 \tilde{\mathcal{R}}(a), \tag{1.194}$$

where $s = Q^2 = -q^2 > 0$ is the physical time-like Minkowski squared momentum transfer. $\tilde{\mathcal{R}}(a)$ are light-by-light corrections which again we neglect, and $\mathcal{R}(a)$ takes

the canonical perturbative form,

$$\mathcal{R}(a) = a + \sum_{n=1}^{\infty} r_n a^{n+1}. \tag{1.195}$$

The $R_{e^+e^-}$ ratio can be related to the Adler D-function using analytic continuation and the optical theorem. Using the optical theorem, we can relate $R_{e^+e^-}$ to the vacuum polarization function,

$$R(s) = \frac{3}{4\pi}\text{Im}\,\Pi(s+i\epsilon) = \frac{3}{8\pi i}[\Pi(s+i\epsilon) - \Pi(s-i\epsilon)]. \tag{1.196}$$

Equation (1.196) can then be related to the Adler D-function by integrating Eq. (1.185) between the limits $s \pm i\epsilon$,

$$\mathcal{R}(s) = \frac{1}{2\pi i}\int_{-s-i\epsilon}^{-s+i\epsilon} dt\,\frac{\mathcal{D}(t)}{t}. \tag{1.197}$$

The above expression can be used to relate d_n to r_n. This expression can also be used to obtain 'experimental' values for the Adler D-function from data for the $R_{e^+e^-}$ ratio [47,48].

1.11 Summary

In this chapter we have introduced some of the basic QCD concepts which will be of use to us in later chapters. We saw how the techniques of path integral quantization and renormalization can be used to obtain perturbative expressions for QCD observables. However, we also saw how the renormalization procedure introduces an element of ambiguity, known as RS dependence to these theoretical predictions. We introduced the beta-function equation and showed how it can be used to determine the scale dependence and also the RS dependence of observables. We then discussed three proposed methods for dealing with RS dependence. We briefly discussed the nature of non-perturbative effects before finally defining some of the observables which we will study in this thesis.

Chapter 2

Deep inelastic scattering

In this chapter we present a brief introduction to the subject of deep inelastic lepton-hadron scattering. Our aim is to show the origin of some of the deep inelastic scattering (DIS) observables we will study in later chapters; in particular, we highlight the role of factorization in computing QCD corrections to these observables. Detailed introductions to the subjects covered in this chapter may be found in Refs. [1–5, 54, 55].

2.1 Introduction

The confinement of quarks within hadrons is a strictly non-perturbative effect – a consequence of which is the absence of free quarks in nature. This apparent unavailability of free quarks with which to test QCD would seem to make any perturbative treatment impossible. Fortunately, the twin concepts of asymptotic freedom and *factorization* provide us with an escape route from this apparent impasse.

As a consequence of asymptotic freedom, the behaviour of 'free' (high energy) quarks is relatively well understood. It is how this behaviour transfers over to the scattering of hadron-bound quarks that causes problems. Fortunately, in certain circumstances, it is possible to 'factor out' the non-perturbative effects such that hadron scattering processes can be described in terms of perturbative results for quark scattering, supplemented by a set of non-perturbative parameters. These parameters are presumably calculable in some non-perturbative scheme. However, because of factorization, they can be treated just as free parameters of the theory; their values obtainable by optimiz-

ing fits to data. In this way QCD can be used to make predictions for hadron scattering and, conversely, hadron scattering can be used to test QCD. However, for our ignorance of non-perturbative effects, we pay a price. The factorization process is a type of renormalization and consequently, freedom in the choice of *factorization scheme* results in a theoretical ambiguity – in analogy with RS dependence. The study of the dependence of DIS observables on the choice of renormalization and factorization schemes forms the basis of much of the research carried out in this thesis.

The material in this chapter is organized as follows:

First, we introduce the DIS lepton-hadron scattering diagram and explain how the naïve parton model can be used to describe this process in terms of structure functions. These structure functions have relevance beyond the naïve parton model and are used to characterize the momentum structure of the constituents of hadrons. As such, they are of great importance in DIS. We describe how the operator product expansion can be used to link structure functions to perturbative QCD, via the definition of Mellin moments of the structure functions. In this way we can compute QCD corrections to results obtained from the naïve parton model. We then introduce the renormalization group equation (RGE) for inserted Green's functions and describe how it can be used to determine the scale dependence of the moments. Next, we describe how factorization allows perturbative and non-perturbative effects to be separated in DIS cross sections, and we draw analogies between the renormalization and factorization procedures. We then define several of the DIS observables which are of interest to us in this thesis.

2.2 The basics of deep inelastic scattering

We are interested in deep inelastic lepton-hadron scattering, as described by the following equation,

$$l(k) + h(p) \;\rightarrow\; l'(k') + X, \tag{2.1}$$

and characterized by the diagram in Fig. 2.1. As the title 'DIS' suggests, we are interested in the case where the momentum transfer from the lepton to the hadron is large, and hence the scattering is inelastic. In Eq. (2.1), X represents the collection

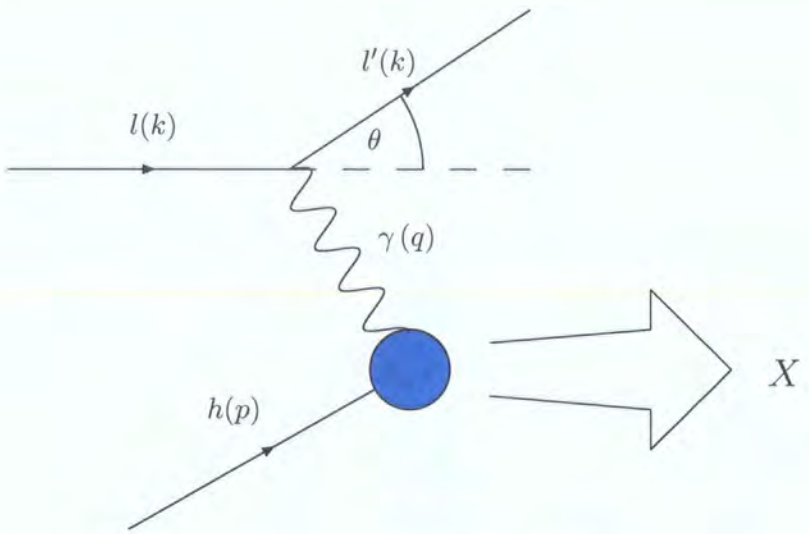


Figure 2.1: The diagram describing deep inelastic scattering.

of final state hadrons; we make no specific assumptions about their nature except that $X \neq h$ (due to the inelastic nature of the scattering). In this thesis we limit ourselves to the case where h is either a proton or neutron, but the results are easily adapted so that they apply to other hadrons.

The final state lepton, l', may be the same as that of the initial state, l, but we also allow for the possibility that it is not. In the case of electron or muon scattering, we have $l = l' = e$ or μ and in these cases, the exchanged vector boson in Fig. 2.1 is a photon.[1] In the case of neutrino scattering, we have $l = \nu$ (or $\overline{\nu}$), whilst l' could be either the original neutrino or its associated lepton. As a result, the scattering could be via the exchange of either a Z or W particle. In effect, we are using an electromagnetic or weak current (via the exchange of a gauge boson) to probe the internal structure of nucleons.

The kinematic variables which characterize DIS are defined below,

$$Q^2 = -q^2, \qquad p^2 = m_N^2, \tag{2.2}$$

$$\nu = p \cdot q, \qquad x = -\frac{q^2}{2p \cdot q} = \frac{Q^2}{2\nu}, \tag{2.3}$$

[1]Here, we make the assumption that the centre of mass energy squared of the scattering is $\ll M_Z^2$.

$$y = \frac{q \cdot p}{k \cdot p}. \tag{2.4}$$

Here, m_N is the mass of the target nucleon. As stated above, the intermediate gauge boson is used to probe the internal structure of the nucleon, and for this to be the case, the initial lepton must transfer a large amount of momentum to it. The region in which the scattering is deeply inelastic and to which we can apply asymptotic freedom, is known as the Bjorken limit. It is defined as,

$$q^2 \rightarrow -\infty, \tag{2.5}$$

$$p \cdot q \rightarrow \infty, \tag{2.6}$$

$$x = \text{constant}. \tag{2.7}$$

The lepton-hadron cross section

The differential cross section for the process in Eq. (2.1) can be written in terms of three independent functions known as structure functions, denoted by F_i, $i = 1, 2, 3$. These functions parameterize the cross section as follows,[2]

$$\frac{d^2\sigma}{dx\,dy} \propto \left[y^2 x F_1 + (1-y)\,F_2 \pm y\left(1 - \frac{y}{2}\right) x F_3\right] \tag{2.8}$$

$$\propto \left[2F_1 \sin^2\frac{\theta}{2} + \frac{m_N^2}{\nu} F_2 \cos^2\frac{\theta}{2} \pm \frac{m_N(k_0 + k_0')}{x\nu} x F_3\right]. \tag{2.9}$$

The value of the constant of proportionality in Eq. (2.8) depends on the type of scattering. Also, the F_3 term is present only in neutrino scattering and the sign in front of it depends on whether l is a neutrino $(+)$ or an anti-neutrino $(-)$.

Using Eq. (2.9), experimental values for the structure functions can be obtained from measurements of the lepton-hadron cross section. At this stage we know nothing about the nature of the structure functions, except that the cross section can be parameterized in terms of them. However, they are physical observables, and so by forging a path from the QCD Lagrangian to the cross section we can test the validity of QCD in describing strong interaction phenomena.

The structure functions describe the internal momentum structure of the nucleon. One

[2]From this point on, we neglect the mass of the nucleon ($Q^2 \gg m_N^2$).

would expect them to be functions of both Q^2 and x, yet they appear to obey the following relation,

$$F_i(x, Q^2) \approx F_i(x). \tag{2.10}$$

This behaviour is known as Bjorken scaling and implies that the constituents of hadrons are point-like. It can be understood in terms of the *naïve parton model* [56–58], of which Eq. (2.10) is a direct prediction. Perturbative QCD corrections to the parton model generate corrections to this in the form of scale logarithms which violate Bjorken scaling. The study of the nature and magnitude of these violations forms the basis of a stringent test of QCD.

For the remainder of this chapter we will discuss DIS in terms of electron or muon scattering, in which the lepton vertex is electromagnetic. We will state, where necessary, how to generalize the results we obtain, to neutrino scattering.

The naïve parton model

In the naïve parton model we consider the constituents of nucleons (which we call partons) to behave as free particles, interacting with the photon probe only. Furthermore, we assume that during its transit across the nucleon, the lepton interacts with only a single parton. These assumptions can be justified by the following arguments:

In the centre of mass frame, and the $Q^2 \to \infty$ limit, the nucleon experiences extreme length contraction and time dilation. Consequently, the lepton 'sees' the nucleon as being flat in its direction of motion and the partons as effectively frozen in time. By the uncertainty principle, the lepton must pass within an $\mathcal{O}(1/Q^2)$ radius of a parton to encounter (interact with) it, and hence the probability of the lepton interacting with more than one parton is suppressed by a factor of,

$$\frac{1/Q^2}{\pi R^2}, \tag{2.11}$$

with R the radius of the nucleon. As a result, it is unlikely that the lepton will interact with more than one parton.

Under these assumptions, the cross section for lepton-hadron scattering can be deter-

mined from two pieces of information: the cross section for lepton-parton scattering and the momentum carried by the individual partons in the nucleon. In light of this, we consider the cross section for a lepton scattering off a quark carrying momentum ξp_μ,

$$l(k) + q(\xi p) \rightarrow l'(k') + q(p'). \tag{2.12}$$

Here, ξ is known as the momentum fraction – it is the fraction of total nucleon momentum being carried by the struck quark. The Feynman rules of section 1.4 can be used to deduce the tree level cross section for the scattering process in Eq. (2.12). It is found to have the form,

$$\frac{d\hat{\sigma}}{d\hat{t}} \propto e_q^2 \frac{\hat{s}^2 + \hat{u}^2}{\hat{s}^2 \hat{t}^2}. \tag{2.13}$$

The hat over the cross section and other variables denotes that we are calculating the cross section in quark (parton) rather than nucleon scattering. The *Mandelstam variables*, $\hat{s}$, $\hat{t}$ and $\hat{u}$ are defined as follows,

$$\hat{s} \equiv (k + \xi p)^2 = \xi \frac{Q^2}{xy}, \tag{2.14}$$

$$\hat{t} \equiv (k - k')^2 = -Q^2, \tag{2.15}$$

$$\hat{u} \equiv (\xi p - k')^2 = \hat{s}(y - 1), \tag{2.16}$$

where we have taken the $Q^2 \rightarrow \infty$ limit.

Equation (2.13) can be converted into a double differential cross section, for the purposes of comparison with Eq. (2.8),

$$\frac{d^2\sigma}{dxdy} \propto [1 + (1 - y)^2] \frac{1}{2} e_q^2 x \delta(x - \xi), \tag{2.17}$$

where we have inserted a factor of $1 = \int_0^1 dx \delta(x - \xi)$, a relation that can be verified by imposing the mass-shell constraint on the outgoing quark (i.e. $k'^2 = 0$).

Comparing Eq. (2.17) with Eq. (2.8), we can make the following identification,

$$\hat{F}_2 = x e_q^2 \delta(x - \xi) = 2x\hat{F}_1. \tag{2.18}$$

This relation is known as the *Callan-Gross* relation and verifies the Bjorken scaling relation of Eq. (2.10), within the naïve parton model. It also suggests that F_2 probes a quark carrying momentum fraction $\xi = x$.

We now come to the central assumption of the parton model. We assume that the quarks in a nucleon can be described by probability density functions, $q(\xi)$, which are representative of the probability of a quark q having momentum ξp_μ. The nucleon structure functions can then be obtained by averaging the quark structure functions, $\hat{F}_k$, over the range of quark momentum fraction $0 \leq \xi \leq 1$, and then summing over all quarks in the nucleon. This average takes the form of the following integral,

$$\begin{aligned} F_2(x) &= \sum_{q,\,\bar{q}} \int_0^1 d\xi q(\xi) \hat{F}_2 \\ &= \sum_{q,\,\bar{q}} e_q^2 x q(x). \end{aligned} \tag{2.19}$$

Intuitively, this makes perfect sense. The momentum structure of the nucleon is simply a product of the combined momentum structure of its constituents. In general, there are contributions to the nucleon structure functions not just from the three constituent quarks (known as valence quarks) but also 'sea quarks' (comprised of virtual quark-antiquark pairs), as well as gluons. Also, beyond the naïve parton model, these results are subject to QCD corrections.

By applying similar reasoning to that used to obtain Eq. (2.19) we can derive nucleon structure functions, in terms of individual quark distribution functions, for a variety of scattering processes. Below we state some of these results which are relevant to later chapters. For simplicity, we consider contributions from the first four quark flavours only. For electron-proton scattering we have the following,

$$F_2^{ep} = \frac{4}{9} x \left[u + \bar{u} + c + \bar{c}\right] + \frac{1}{9} x \left[d + \bar{d} + s + \bar{s}\right], \tag{2.20}$$

and in (anti)neutrino-proton scattering we can obtain,

$$F_2^{\nu p} = 2x \left[d + s + \bar{u} + \bar{c}\right], \tag{2.21}$$

$$F_2^{\bar{\nu} p} = 2x \left[u + c + \bar{d} + \bar{s}\right]. \tag{2.22}$$

From these, F_1 structure functions can be obtained using the Callan-Gross relation, $2xF_1 = F_2$.

We can also define a set of spin-dependent structure functions, g_1 and g_2, which describe the scattering of beams of leptons and hadrons polarized either parallel or anti-parallel to each other. These structure functions characterize the spin structure of nucleons and from them, several useful observables can be derived. In the naïve parton model they are given by,

$$g_1(x) = \frac{1}{2}\sum_q e_q^2 \left[\Delta q(x) + \Delta\bar{q}(x)\right], \tag{2.23}$$

$$g_2(x) = 0, \tag{2.24}$$

where $\Delta q = q^\uparrow - q^\downarrow$. The functions $q^\uparrow$ and $q^\downarrow$ are distribution functions representing the probability that the spin of the quark is respectively parallel and anti-parallel to the spin of its parent nucleon. The above result for g_2 is misleading and when a more sophisticated version of the parton model is applied, g_2 has a magnitude similar to that of g_1 (see Ref. [59]).

The naïve parton model gives us an intuitive grasp of the behaviour of nucleon constituents. However, as the title suggests, this view is rather over-simplistic and it is easy to imagine, qualitatively, how Bjorken scaling is violated by radiative corrections. In order to quantify this violation, it is necessary to perform a more rigorous analysis of DIS, within QCD.

2.3 Field theory description of DIS

If we assume that the scattering in Fig. 2.1 occurs via photon exchange then the amplitude for the process can be written as,

$$\mathcal{A} = e\overline{u}(k')\gamma^\mu u(k)\frac{1}{q}^2 \langle X|j_\mu(0)|N,p\rangle, \tag{2.25}$$

where the $u(k)$'s represent spinors for the incoming and outgoing leptons and $j_\mu(x)$ is the electromagnetic current. The cross section is proportional to the amplitude squared, and because of the nature of the diagram in Fig. 2.1, we can separate it out

into leptonic and hadronic components $L_{\mu\nu}$ and $W_{\mu\nu}$,

$$\frac{d^2\sigma}{dxdy} \propto L_{\mu\nu}W^{\mu\nu}. \tag{2.26}$$

For photon exchange, $L_{\mu\nu}$ is fairly simple and describes the nature of the virtual photon probe. At leading order (LO) it has the form,

$$L_{\mu\nu} = e^2 Tr[\not{k}\gamma_\mu \not{k}\gamma_\nu]. \tag{2.27}$$

The hadron vertex in Fig. 2.1 dominates the energy dependence of DIS observables and because of this, we neglect radiative corrections to the lepton vertex.

$W_{\mu\nu}$ is descriptive of the internal structure of the nucleon and from Eqs. (2.25) and (2.26) can be determined to be,

$$W_{\mu\nu}(p,q) = \frac{1}{2\pi}\int d^4z\, e^{iq\cdot z}\langle N,p|[j_\mu(z),j_\nu(0)]|N,p\rangle. \tag{2.28}$$

We can project out different tensor structures of $W_{\mu\nu}$, and parameterize these structures using three functions: W_i $(i = 1,2,3)$,

$$W_{\mu\nu}(p,q) = \left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{q^2}\right)W_1 + \left(p_\mu - \frac{p\cdot q}{q^2}q_\mu\right)\left(p_\nu - \frac{p\cdot q}{q^2}q_\nu\right)W_2. \tag{2.29}$$

In the case of neutrino scattering, Eq. (2.29) must be supplemented by an additional term,

$$-i\varepsilon_{\mu\nu\alpha\beta}\frac{p^\alpha q^\beta}{\nu}\nu W_3. \tag{2.30}$$

Via the optical theorem, $W_{\mu\nu}$ can be related to the forward scattering amplitude for $\gamma N \to \gamma N$, given by,

$$T_{\mu\nu}(p,q) = i\int d^4z e^{iq\cdot z}\langle N,p|T[j_\mu(z)j_\nu(0)]|N,p\rangle \tag{2.31}$$

$$= \left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{q^2}\right)T_1 + \left(p_\mu - \frac{p\cdot q}{q^2}q_\mu\right)\left(p_\nu - \frac{p\cdot q}{q^2}q_\nu\right)T_2, \tag{2.32}$$

where

$$W_i = \frac{1}{\pi}\mathrm{Im}\, T_i. \tag{2.33}$$

The W_i functions are directly related to the structure functions of Eq. (2.8) by,

$$F_L(x,Q^2) \equiv F_2 - 2xF_1 = \nu W_2 - 2xW_1, \tag{2.34}$$

$$F_2(x,Q^2) = \nu W_2, \tag{2.35}$$

$$F_3(x,Q^2) = \nu W_3. \tag{2.36}$$

From Eq. (2.31), it is clear that the product of operators $T[j_\mu(z)j_\nu(0)]$ is our link from the structure functions to the QCD Lagrangian. The study of this object and the extraction from it of perturbative predictions forms the basis of the next few sections.

2.4 The operator product expansion approach to DIS

Our link between the hadronic tensor and the QCD Lagrangian is the time-ordered product of operators $T[j_\mu(x)j_\nu(0)]$. The OPE, which we discussed very briefly in chapter 1, allows us to expand this product in terms of a set of well-defined operators, $\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}}(x)$,

$$iT[j_\mu(x)j_\nu(0)] = \sum_i C^i_{\mu\nu\mu_1\cdots\mu_{l_i}}(x^2)\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}}(0), \tag{2.37}$$

where the expansion is in powers x^2 (*light-cone* expansion). The most dominant terms in Eq. (2.37) are those for which the Wilson coefficients, $C_i(x^2)$ are most singular, and therefore our task must be to identify these terms.

From the nature of the current $j_\mu(x)$, we can determine the appropriate set of operators in Eq. (2.37) [60, 61] (see also Refs. [62–64]). We identify three different types of operator: a set of non-singlet[3] (NS) operators,

$$\mathcal{O}^{\mu_i\cdots\mu_{l_i}}_{a,\,\mathrm{NS}\pm} = \frac{i^{l_i-1}}{l!}(\overline{\psi}\lambda_a\gamma^{\mu_1}(1\pm\alpha\gamma_5)D^{\mu_2}\ldots D^{\mu_{l_i}}\psi + \text{permutations}) - \text{trace terms}, \tag{2.38}$$

and two sets of singlet operators associated respectively with quark and gluon fields,

$$\mathcal{O}^{\mu_i\cdots\mu_{l_i}}_{q\pm} = \frac{i^{l_i-1}}{l!}(\overline{\psi}\gamma^{\mu_1}(1\pm\alpha\gamma_5)D^{\mu_2}\ldots D^{\mu_{l_i}}\psi + \text{permutations}) - \text{trace terms},$$

[3]Here the terms 'non-singlet' and 'singlet' refer to SU(N_f) of flavour.

(2.39)

$$\mathcal{O}_g^{\mu_i\ldots\mu_{l_i}} = \frac{i^{l_i-2}}{(2l)!}(F_{a_1}^{\mu\mu_1} D_{a_1a_2}^{\mu_2}\ldots D_{a_{l_i-2}a_{l_i-1}}^{\mu_{l_i-1}}(F_{\mu}^{\mu_{l_i}})_{a_{l_i}} + \text{permutations}) - \text{trace terms}.$$

(2.40)

Here, the addition of permutations has the effect of symmetrizing the operators in their Lorentz indices, and λ_a are the generators of SU(N_f). The value of α depends on the type of scattering considered. For neutral current scattering with $Q^2 \ll M_Z^2$ we have $\alpha = 0$, and for charged current scattering $\alpha = 1$.

The Wilson coefficients can be decomposed into contributions from the three different tensor structures that will contribute to each of the T_i functions,

$$\begin{aligned} C^i_{\mu\nu\mu_1\ldots\mu_{l_i}}(x^2) &= -h_i^{(1)}(x^2) g_{\mu\nu} x_{\mu_1}\ldots x_{\mu_{l_i}} \\ &\quad + h_i^{(2)}(x^2) g_{\mu\mu_1} g_{\nu\mu_2} x_{\mu_3}\ldots x_{\mu_{l_i}} \\ &\quad - h_i^{(3)}(x^2)\varepsilon_{\mu\nu\mu_1\alpha} x^\alpha x_{\mu_2}\ldots x_{\mu_{l_i}} + \ldots . \end{aligned} \tag{2.41}$$

In the absence of dimensionful parameters, the functions $h_i^{(k)}(x^2)$ consist of only a single power of x^2. Consequently, the dominant operator contributions in Eq. (2.37) are those for which the value of this power is lowest. By applying dimensional analysis to Eqs. (2.37) and (2.41), we can determine that these dominant operators are those which have lowest *twist*, t_i, defined as,

$$t_i \equiv d_i - l_i, \tag{2.42}$$

where d_i and l_i are, respectively, the mass dimension and spin of the operators $\mathcal{O}_i^{\mu_1\ldots\mu_{l_i}}$. The operators in Eqs. (2.38) - (2.40) have t_i =2 and represent the lowest twist operators contributing to Eq. (2.37); we neglect all operators that are of twist-4 or higher.

The Wilson coefficients of Eq. (2.41) can be rewritten in terms of partial derivative tensors,

$$\begin{aligned} C^i_{\mu\nu\mu_1\ldots\mu_{l_i}}(x^2) &= -(-1)^{l_i} g_{\mu\nu}\partial_{\mu_1}\partial_{\mu_2}\partial_{\mu_3}\ldots\partial_{\mu_{l_i}} H_i^{(1)}(x^2) \\ &\quad + (-1)^{l_i-2} g_{\mu\mu_1} g_{\nu\mu_2}\partial_{\mu_3}\ldots\partial_{\mu_{l_i}} H_i^{(2)}(x^2) \\ &\quad - (-1)^{l_i}\varepsilon_{\mu\nu\mu_1\alpha}\partial^\alpha\partial_{\mu_2}\partial_{\mu_3}\ldots\partial_{\mu_{l_i}} H_i^{(3)}(x^2), \end{aligned} \tag{2.43}$$

where the functions $H_i^{(k)}(x^2)$ are constructed such that they reproduce $h_i^{(k)}(x^2)$ when acted upon by the derivatives. The three components of Eq. (2.43) correspond to the three different tensor structures of Eq. (2.32). Therefore, each of the functions $H_i^{(1)}(x^2)$, $H_i^{(2)}(x^2)$ and $H_i^{(3)}(x^2)$ is relevant to a particular structure function. This can be seen by substituting Eqs. (2.43) and (2.37) into Eq. (2.31) yielding,

$$\begin{aligned} T_{\mu\nu}(p,q) &= \sum_i \mathcal{A}_i \left(-g_{\mu\nu}(2x)^{-l_i} \tilde{H}_i^{(1)}\left(q^2\right) + \frac{p_\mu p_\nu}{p\cdot q}(2x)^{-l_i+1} \tilde{H}_i^{(2)}\left(q^2\right) \right. \\ &- \left. \varepsilon_{\mu\nu\alpha\beta}\frac{p^\alpha q^\beta}{p\cdot q}(2x)^{-l_i} \tilde{H}_i^{(3)}\left(q^2\right) \right), \end{aligned} \tag{2.44}$$

where $\mathcal{A}_i$ is a set of constants derived from the nucleon matrix element,

$$\langle N,p|\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}}|N,p\rangle = \mathcal{A}_i p^{\mu_1}\ldots p^{\mu_{l_i}} + \ldots, \tag{2.45}$$

and the functions $\tilde{H}_i^{(k)}\left(q^2\right)$ are constructed from the Fourier transforms of $H_i^{(k)}(x^2)$,

$$\tilde{H}_i^{(k)}(q^2) = (-q^2)^{l_i-\delta_k}\int d^4x e^{iq\cdot x} H_i^{(k)}(x), \tag{2.46}$$

where $\delta_1 = \delta_3 = 0$ and $\delta_2 = 1$. Comparing Eq. (2.44) with Eq. (2.32) we can now make the following identifications,

$$T_1(p,q) = \sum_i \mathcal{A}_i(2x)^{-l_i}\tilde{H}_i^{(1)}(q^2), \tag{2.47}$$

$$\nu T_2(p,q) = \sum_i \mathcal{A}_i(2x)^{-l_i+1}\tilde{H}_i^{(2)}(q^2), \tag{2.48}$$

$$\nu T_3(p,q) = \sum_i \mathcal{A}_i(2x)^{-l_i}\tilde{H}_i^{(3)}(q^2). \tag{2.49}$$

The next step is to show how the functions $\tilde{H}_i^{(k)}(q^2)$ can be related to the structure functions via their Mellin moments.

Moments of structure functions

Using the OPE approach, we cannot predict the energy dependence of the structure functions directly, rather, we can do so only indirectly via the moments of the structure functions. We define the moments as the following Mellin transforms of F_i,

$$\mathcal{M}_1(n;Q^2) = \int_0^1 dx\, x^{n-1}F_1(x,Q^2), \qquad \mathcal{M}_2(n;Q^2) = \int_0^1 dx\, x^{n-2}F_2(x,Q^2),$$

$$\mathcal{M}_3(n;Q^2) = \int_0^1 dx\, x^{n-1}F_3(x,Q^2). \tag{2.50}$$

If we assume that the functions T_i are symmetric under the transformation,[4] $x \to -x$, then we can infer from Eqs. (2.47) - (2.49) that they will have a branch cut running from $-1 \leq x \leq 1$. Using Eqs. (2.34) - (2.36) together with Eq. (2.33), and the existence of the branch cut discontinuity, $\mathcal{M}_k(n;Q^2)$ can be written in terms of the functions T_i,

$$\mathcal{M}_1(n;Q^2) = \frac{1}{4\pi i}\int_{\mathscr{C}} dx x^{n-1}T_1, \qquad \mathcal{M}_2(n;Q^2) = \frac{1}{4\pi i}\int_{\mathscr{C}} dx x^{n-2}\nu T_2, \tag{2.51}$$

$$\mathcal{M}_3(n;Q^2) = \frac{1}{4\pi i}\int_{\mathscr{C}} dx x^{n-1}\nu T_3, \tag{2.52}$$

where $\mathscr{C}$ represents a large, circular, anticlockwise contour in the complex x-plane. This contour integration has the effect of isolating individual terms in the Laurent expansions of Eqs. (2.47) - (2.49); via the residue theorem, this leads to,

$$\mathcal{M}_k(n;Q^2) = \frac{2^{\delta_k}}{2^{n+2}}\sum_{i,l_i=n} \mathcal{A}_i\, \tilde{H}_i^{(k)}(q^2), \tag{2.53}$$

where, for the nth moment, the sum is over all Wilson coefficients[5] $\tilde{H}_i^{(k)}(q^2)$ corresponding to operators of spin $l_i = n$.

The key to linking the Wilson coefficients to QCD lies in studying the RG behaviour of the composite operators $\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}}$. Through this study, it is possible to predict the scale dependence of $\tilde{H}_i^{(k)}(q^2)$, and hence of the moments, giving us access to a rich testing ground for QCD phenomenology.

[4] The following derivation is slightly more troublesome in the case of F_3. However, the final result holds for all structure functions.

[5] Although $C_i(x^2)$ are the true Wilson coefficients, we also refer to $H_i^{(k)}(x^2)$ and $\tilde{H}_i^{(k)}(q^2)$ as such, due to their close connection.

2.5 Renormalization group analysis of composite operators

In this section, we study the RG behaviour of the composite operators in Eqs. (2.38) - (2.40), with the intent of determining the scale dependence of the moments. This analysis runs parallel to that carried out in section 1.5 and follows the same line of reasoning.

The operators of Eqs. (2.38) - (2.40) become singular upon the inclusion of radiative corrections. These singularities are different in nature to those considered in section 1.5.1; they are caused by low energy (soft), non-perturbative QCD interactions which cannot be treated perturbatively. The singularities can be dealt with using renormalization, but this particular form of renormalization is termed factorization, in reference to its association with the 'factoring out' of non-perturbative physics. As a consequence of this new renormalization, the operators $\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}}$ depend on the relevant renormalization scale and the subtraction procedure used to define the finite part of the regularization. Together, these define the factorization scheme (FS).

To study the behaviour of $\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}}$ under variations in RS parameters, we consider the behaviour, under renormalization, of OPI Green's functions constructed from ordinary QCD fields *and* composite operators. By this we mean Green's functions in which one or more external legs/fields are replaced by an operator of the type in Eqs. (2.38) - (2.40). Figure 2.2 shows the Feynman diagram depiction for such a Green's function; the double line represents the inserted operator. Green's functions constructed from composite operators are termed 'inserted' and are denoted by using a subscript $\mathcal{O}_i$. Just as ordinary Green's functions require renormalization, so too do inserted Green's functions, due to the above-mentioned singularities in the bare operators of Eqs. (2.38) - (2.40).

In analogy with Eq. (1.86) we can relate the renormalized inserted Green's functions to their bare counterparts,

$$\tilde{G}^{(n)}_{\mathcal{O}_i}(\mathbf{p},\bar{g},\xi,m,M) \;=\; Z_{ij}Z_A^{n_G/2}Z_\psi^{n_\psi/2}\tilde{G}^{(n)}_{\mathcal{O}_{j,B}}(\mathbf{p},g_B,\xi_B,m_B). \tag{2.54}$$

Here, Z_A and Z_ψ are the renormalization constants from Eq. (1.72), and M is a new renormalization scale, called the factorization scale. M is distinct from μ, it con-

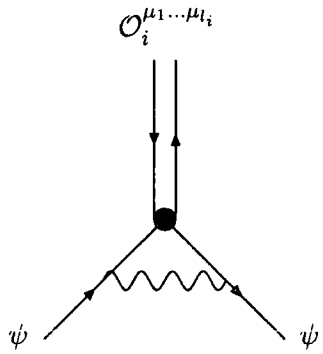


Figure 2.2: Example of a diagram for an inserted Green's function. The double line represents the insertion of a composite field replacing a normal external propagator.

trols the renormalization of composite operators and determines at what scale non-perturbative effects are factored out. Also, we have limited ourselves to inserted Green's functions with only one operator insertion, although the results generalize easily.

In Eq. (2.54) we have introduced a new renormalization constant, Z_{ij}. This constant controls the renormalization of the operators, serving to remove singularities in $\mathcal{O}_i$ generated by radiative corrections. It is defined (in analogy with Eq. (1.72)) by,

$$(\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}})_B = Z_{ij}\mathcal{O}_j^{\mu_1\cdots\mu_{l_j}}, \tag{2.55}$$

where $\mathcal{O}_i^{\mu_1\cdots\mu_{l_i}}$ written without a subscript B now denotes the renormalized operator. For the remainder of this section (and as we have done already in Eq. (2.54)) we omit the string of Lorentz indices from operators, Green's functions and coefficient functions.

In general, Z_{ij} is a matrix. For instance, in the case of singlet operator contributions, Eq. (2.55) is of the form,

$$(\mathcal{O}_i^{\rm S})_B = Z_{ij}^{\rm S}\mathcal{O}_j^{\rm S}, \qquad i,j = q,\, g. \tag{2.56}$$

However, in the NS case, Eq. (2.55) simplifies to,

$$(\mathcal{O}^{\rm NS})_B = Z^{\rm NS}\mathcal{O}^{\rm NS}, \tag{2.57}$$

where $Z^{\rm NS}$ is not a matrix.

The fact that the bare Green's functions must be independent of RG parameters allows us to derive the RGE for inserted Green's functions. In analogy with Eq. (1.87), we have,

$$\left[\delta_{ij}\left(M\frac{\partial}{\partial M}+\beta_{\bar{g}}\frac{\partial}{\partial \bar{g}}+\mathbb{D}\right)-\gamma_{ij}\right]\tilde{G}^{(n)}_{\mathcal{O}_i}(\mathbf{p},\bar{g},\xi,m,M) \;=\; 0, \tag{2.58}$$

where $\mathbb{D}$ is defined in Eq. (1.89) and we have defined an object known as the operator anomalous dimension matrix, γ_{ij}; it is related to the renormalization constants Z_{ij} by,

$$\gamma_{ij}Z_{jk} \;=\; -M\frac{\partial Z_{ik}}{\partial M}. \tag{2.59}$$

The above renormalization constant depends on n and therefore, so does γ_{ij}. By multiplying the expression in Eq. (2.37) by n_G gauge fields and n_ψ fermion fields and then taking the Fourier transform, we obtain an equation relating $\tilde{G}_{j_\mu j_\nu}$ (an inserted Green's function formed from the operator product on the LHS of Eq. (2.37)) and $\tilde{G}_{\mathcal{O}_i}$,

$$i\tilde{G}^{(n)}_{j_\mu j_\nu}(q,-q;\mathbf{p}) \;=\; \sum_i C_i(q)\tilde{G}^{(n)}_{\mathcal{O}_i}(\mathbf{p}). \tag{2.60}$$

The current j_μ is conserved and therefore requires no renormalization further to that carried out in section 1.5. Consequently, $\tilde{G}_{j_\mu j_\nu}$ has no operator anomalous dimension and therefore obeys the RGE of Eq. (1.87), with $\mu \to M$. Acting on Eq. (2.60) with the term in the square bracket in Eq. (2.58), and then using Eq. (1.87) gives us an RGE for the coefficient functions C_i,

$$\left[\delta_{ij}\left(M\frac{\partial}{\partial M}+\beta_{\bar{g}}\frac{\partial}{\partial \bar{g}}+\beta_\xi\frac{\partial}{\partial \xi}-\gamma_m m\frac{\partial}{\partial m}\right)+\gamma_{ji}\right]C_j \;=\; 0. \tag{2.61}$$

Substituting for C_i using Eqs. (2.43) and (2.46) then gives,

$$\left[\delta_{ij}\left(M\frac{\partial}{\partial M}+\beta_{\bar{g}}\frac{\partial}{\partial \bar{g}}+\beta_\xi\frac{\partial}{\partial \xi}-\gamma_m m\frac{\partial}{\partial m}\right)+\gamma_{ji}\right]\tilde{H}^{(k)}_j \;=\; 0. \tag{2.62}$$

We can now see how, by solving the above equation, the scale dependence of the moments can be obtained and hence how the structure functions can be compared with QCD predictions over a range of energies. The simplest case is that in which we have contributions from only NS operators. Because there is only one of these operators of

lowest twist per value of spin (and hence of n), we can write the moments as,

$$\mathcal{M}_k(n;Q^2) \;=\; \mathcal{A}_n \tilde{H}_n^{(k)}(Q^2). \tag{2.63}$$

The scale dependence of the moments can then be determined by solving Eq. (2.62), with $\mathcal{A}_n$ being determined by fitting to data.

A brief note on the calculation of the anomalous dimension

The anomalous dimension of Eq. (2.59) describes the renormalization of inserted Green's functions, and hence its perturbative expansion can be computed by considering loop corrections to these functions. In the case of NS operators, the anomalous dimension is particularly simple. We need to consider the inserted Green's function constructed from two external quark legs and an NS, twist two, spin n operator of the type in Eq. (2.38). From Eq. (2.54) we can determine that this undergoes the following renormalization,

$$\tilde{G}^{(2)}_{\mathcal{O}_n^{\rm NS}} \;=\; Z_n^{\rm NS} Z_\psi \left(\tilde{G}^{(2)}_{\mathcal{O}_n^{\rm NS}}\right)_B. \tag{2.64}$$

In analogy with Eq. (1.76) and Fig. 1.8, this renormalization can be represented diagrammatically in terms of diagrams of the type shown in Fig. 2.2. The evaluation of these diagrams is slightly more difficult than that of ordinary Green's functions, but the principles are the same: regularization of divergent integrals followed by determination of the singular nature of the renormalization constants. The calculation was first performed in Ref. [65], and from the results therein, we can determine that $Z_n^{\rm NS}$ has the form,

$$Z_n^{\rm NS} \;=\; 1 + \frac{a}{4\varepsilon} C_F \left(1 - \frac{2}{n(n+1)} + 4\sum_{s=2}^{n} \frac{1}{s}\right). \tag{2.65}$$

From this, and using Eq. (2.59), we can evaluate the anomalous dimension to order a,

$$\gamma_{{\rm NS},n}(a) \;=\; -\frac{1}{Z_n^{\rm NS}} \frac{\partial}{\partial \ln M} Z_n^{\rm NS} \tag{2.66}$$

$$\;=\; -\frac{a}{2} C_F \left(1 - \frac{2}{n(n+1)} + 4\sum_{s=2}^{n} \frac{1}{s}\right). \tag{2.67}$$

In general, the anomalous dimension takes the form of a perturbative expansion;

$$\gamma_{\mathrm{NS},n}(a) = -d(n)a - d_1(n)a^2 - d_2(n)a^3 - \ldots, \tag{2.68}$$

and from Eq. (2.67), we can determine $d(n)$ to be,

$$d(n) = \frac{C_F}{2}\left(1 - \frac{2}{n(n+1)} + 4\sum_{s=2}^{n}\frac{1}{s}\right). \tag{2.69}$$

In the singlet case, the anomalous dimension is a matrix and has the form,

$$\gamma_{\mathrm{S},n}(a) = \begin{pmatrix} \gamma^{qq}(a) & \gamma^{qg}(a) \\ \gamma^{gq}(a) & \gamma^{gg}(a) \end{pmatrix} \tag{2.70}$$

$$= -\boldsymbol{d}(n)a - \boldsymbol{d}_1(n)a^2 - \boldsymbol{d}_2(n)a^3 - \ldots. \tag{2.71}$$

The coefficient matrix $\boldsymbol{d}$ was first obtained in Ref. [60], the elements have the following form,

$$d^{qq}(n) = \frac{C_F}{2}\left(1 - \frac{2}{n(n+1)} + 4\sum_{s=2}^{n}\frac{1}{s}\right), \tag{2.72}$$

$$d^{qg}(n) = -\frac{2(n^2+n+2)}{n(n+1)(n+2)}T_R N_f, \tag{2.73}$$

$$d^{gq}(n) = -C_F\frac{n^2+n+2}{n(n^2-1)}, \tag{2.74}$$

$$d^{gg}(n) = \frac{C_A}{2}\left(\frac{1}{3} - \frac{4}{n(n-1)} - \frac{4}{(n+1)(n+2)} + 4\sum_{s=2}^{n}\frac{1}{s}\right) + \frac{2}{3}T_R N_f. \tag{2.75}$$

The first NLO calculations of the non-singlet and singlet anomalous dimension were carried out in Refs. [66, 67] and [68, 69] respectively. These results were subsequently obtained in either simplified or corrected forms in Refs. [70–74]. NNLO non-singlet and singlet coefficients were first obtained in Refs. [75] and [76] respectively, but only for a limited set[6] of values of n. This set was then extended in Ref. [77], but only recently have analytic expression for the anomalous dimension, applicable to both odd and even n, been obtained [78, 79].

[6]The singlet and non-singlet coefficients were only available for even and odd values of n respectively.

2.6 Factorization

The renormalization carried out on the operators $\mathcal{O}_i$ in section 2.5 is more commonly known as factorization, in reference to the fact that we are separating (or factoring out) long range and short range interactions. Our lack of knowledge of non-perturbative dynamics inside the nucleon manifests itself as singularities in the radiative corrections to the operators in Eqs. (2.38) - (2.40); these are indicative of the fact that we are applying perturbation theory to an energy region where it ought not to be applied. The factorization scale M represents a boundary between long-range and short range effects. Above M we apply perturbation theory, but below, we parameterize the unknown non-perturbative effects using the constants $\mathcal{A}_n$. As a consequence, the range of validity of this approximation mirrors that of the perturbative approximation itself. For large M, we are applying perturbation theory to the higher Q region only and therefore the factorized result will be a good approximation. However, the lower we push M the more we step out of the energy range in which we should be applying perturbation theory, and therefore the factorized result tends to an unphysical limit as $M \to \Lambda$. In this way, the concept of asymptotic freedom propagates through to factorized results.

Of course, as we might have expected, there is a price to pay for our ignorance of the non-perturbative effects. In order to define factorized predictions for observables a factorization scheme (FS) must be chosen. As a result of this freedom of choice, predictions for the moments will now suffer from both factorization *and* renormalization scheme (FRS) dependence [80]. This doubles the level of ambiguity and hence increases the theoretical error of any predictions made. The standard approach to this is to choose some physical scale (such as $M = \mu = Q$) and apply the same scheme (such as $\overline{\text{MS}}$) to both renormalization and factorization. Although this setting of scales is convenient, it is not necessary. We shall show in chapter 6 that we can quantify factorization scheme dependence in the same way we have done with RS dependence, and consequently we can apply the ideas behind CORGI to FRS dependence.

2.7 DIS observables of interest

The DIS observables of interest in this thesis fall into two categories. These are sum rules and moments of structure functions.

2.7.1 Sum rules

Sum rules are combinations of structure functions integrated over the full range of x. Technically they are just combinations of the lowest n moments. However, the combinations are chosen such that they give particularly simple results in the parton model, often related to conserved quantities. These results are then subject to QCD corrections, and therefore the sum rules serve as a good test of both the parton model and QCD.

In the language of the OPE, the reason for the simplicity of sum rules is the existence of certain combinations of operators which do not require renormalization, and hence have vanishing anomalous dimensions. Because of this they do not suffer from FS dependence. Like all DIS observables, the sum rules are Euclidean i.e. they are defined for space-like momenta ($Q^2 \equiv -q^2 > 0$) only.

The Gross-Llewelyn Smith sum rule

The Gross-Llewelyn Smith (GLS) sum rule [81] is derived from (anti)neutrino-proton scattering, and is defined by the following integral,

$$K_{GLS}(Q) = \frac{1}{2}\int_0^1 dx \left(F_3^{\nu p} + F_3^{\bar{\nu} p}\right). \tag{2.76}$$

Using the results from the naïve parton model in Eqs. (2.20) and (2.21), this sum rule can be related to an integral over valence quark distribution functions,

$$\begin{aligned} K_{GLS}(Q) &= \int_0^1 dx \left(d - \bar{u} + u - \bar{d}\right) \\ &= \int_0^1 dx \left(d_V + u_V\right) \\ &= 3, \end{aligned} \tag{2.77}$$

and this implies that $K_{GLS}(Q^2)$ counts the combined number of up and down quarks in the proton. Here, $q_V \equiv q - \bar{q}$ is known as the valence quark distribution and represents the number of 'non-virtual' quarks. The parton model result in Eq. (2.77) is subject to perturbative corrections,

$$K_{GLS} = 3\left(1 - \frac{3}{4}C_F\mathcal{K}(a)\right). \tag{2.78}$$

Here, we have ignored corrections from light-by-light scattering diagrams. The perturbative corrections $\mathcal{K}(a)$ take the canonical form of a series expansion in powers of a,

$$\mathcal{K}(a) = a + \sum_{n=1}^{\infty} k_n a^{n+1}. \tag{2.79}$$

The coefficients k_1 and k_2 were obtained in Refs. [82, 83] and [84] respectively. In $\overline{\text{MS}}$ with $\mu = Q$ they are given by,

$$k_1 = -\frac{1}{3}N_f + \left(\frac{23}{12}C_A - \frac{7}{8}C_F\right), \tag{2.80}$$

$$\begin{aligned} k_2 = {} & N_f^2\left(\frac{115}{648}\right) + N_f\left(-\frac{3535}{1296} - \frac{\zeta_3}{2} + \frac{5}{9}\zeta_5\right)C_A + N_f\left(\frac{133}{864} + \frac{5}{18}\zeta_3\right)C_F \\ & + C_A^2\left(\frac{5437}{648} - \frac{55}{18}\zeta_5\right) + C_A C_F\left(-\frac{1241}{432} + \frac{11}{9}\zeta_3\right) + C_F^2\left(\frac{1}{32}\right). \end{aligned} \tag{2.81}$$

The unpolarized Bjorken sum rule

The unpolarized Bjorken sum rule [85] is also derived from (anti)neutrino-proton scattering, but refers to F_1. It is defined as follows,

$$U_{uBj}(Q) \equiv \int_0^1 dx \left(F_1^{\bar{\nu}p} - F_1^{\nu p}\right) \tag{2.82}$$

$$= 1 - \frac{1}{2}C_F\mathcal{U}(a), \tag{2.83}$$

where,

$$\mathcal{U}(a) = a + \sum_{n=1}^{\infty} u_n a^{n+1}. \tag{2.84}$$

The LO, NLO and NNLO coefficients of $\mathcal{U}(a)$ were obtained in Refs. [86], [87] and [88] respectively. In $\overline{\text{MS}}$ with $\mu = Q$ they have the following form,

$$u_1 = -\frac{4}{9}N_f + \left(\frac{91}{36}C_A - \frac{11}{8}C_F\right), \tag{2.85}$$

$$u_2 = N_f^2\left(\frac{115}{648}\right) + N_f\left(-\frac{4235}{1296} + \frac{7}{6}\zeta_3 - \frac{5}{3}\zeta_5\right)C_A + N_f\left(\frac{335}{288} - \frac{1}{6}\zeta_3\right)C_F$$

$$+C_A^2\left(\frac{8285}{648}+5\zeta_3-10\zeta_5\right)+C_AC_F\left(-\frac{2731}{144}-\frac{91}{3}\zeta_3+\frac{95}{2}\zeta_5\right)$$
$$+C_F^2\left(\frac{313}{32}+\frac{47}{2}\zeta_3-35\zeta_5\right). \tag{2.86}$$

The polarized Bjorken sum rule

The polarized Bjorken sum rule [89, 90] is the polarized scattering analogue of K_{uBj}. This sum rule is based on current algebra and SU(2) isospin symmetry, and is defined as,

$$K_{pBj}(Q) \;=\; \int_0^1 dx\,(g_1^{ep}-g_1^{en})\,. \tag{2.87}$$

As before, this sum rule can be split into a parton model result and perturbative corrections. These corrections are identical to those for K_{GLS} except that K_{pBj} has no corrections due to light-by-light scattering,

$$K_{pBj} \;=\; \frac{1}{6}\left|\frac{g_A}{g_V}\right|\left(1-\frac{3}{4}C_F\mathcal{K}(a)\right). \tag{2.88}$$

Finally, there are a few general points to be made in relation to DIS sum rules:

- Each of the above sum rules will receive non-perturbative corrections of the form,

$$\sum_{n=1}^{\infty}\mathcal{C}_n\left(\frac{\Lambda^2}{Q^2}\right)^n. \tag{2.89}$$

  These correspond to contributions from operators in Eq. (2.37) with twist greater than two and consequently, they are commonly known as higher-twist corrections. They are suppressed relative to the perturbative corrections due to the sub-dominance of non-leading twist contributions in the OPE. In the GLS case the magnitude of these corrections has been estimated in [91–96].

- As a result of the anomalous dimension relevant to each sum rule vanishing, they do not suffer from a dependence on the choice of factorization scheme; consequently we need only consider their renormalization scheme dependence.

- Both the polarized Bjorken and GLS sum rules have been measured in experiment in Refs. [97] and [98–102] respectively. Currently, no data exists for the

unpolarized Bjorken sum rule.

2.7.2 Moments of structure functions

The moments were defined for each of the structure functions in Eq. (2.50). In this thesis, we will be concerned with moments of F_3 in neutrino-nucleon scattering, and moments of F_2 in electron-proton scattering.

Moments of F_3 in neutrino-nucleon scattering

These moments are defined as follows,

$$\mathcal{M}_3^{\nu N}(n;Q^2) \;=\; \int_0^1 x^{n-1} F_3^{\nu N} dx. \tag{2.90}$$

Due to the fact that the only lowest twist operators which contribute to these moments are of NS nature, the expression for these moments is particularly simple,

$$\mathcal{M}_3(n;Q^2) \;=\; \mathcal{A}_n \tilde{H}_n^{(3)}(Q^2). \tag{2.91}$$

From now on, we drop the νN superscript, and assume that the nature of the scattering is understood.

In section 2.5 we saw how the scale dependence of the moments can be obtained in terms of the beta-function equation and anomalous dimension, via the RGE. From this point on, we change notation slightly to that used in Ref. [103]; this facilitates the separate study of factorization and renormalization dependence. The moments can be factorized in the following form,

$$\mathcal{M}_3^{\nu N}(n;Q^2) \;=\; \langle N|\mathcal{O}_{n,\mathrm{NS}}(M)|N\rangle \mathcal{C}_n(Q,M,\mu,a(\mu)), \tag{2.92}$$

where $\langle N|\mathcal{O}_{n,\mathrm{NS}}(M)|N\rangle$ is known as the operator matrix element. It is derived from the operator in Eq. (2.38), factorized at the scale M. $\mathcal{C}_n(Q,M,\mu,a(\mu))$ is the coefficient function; it is derived from the Wilson coefficients in Eq. (2.43) and depends on both the renormalization *and* the factorization scheme.

The M dependence of the NS operator matrix element is governed by its anomalous

dimension,

$$\frac{M}{\langle\mathcal{O}(M)\rangle}\frac{\partial\langle\mathcal{O}(M)\rangle}{\partial M} = \gamma_{n,\mathrm{NS}}(a) = -d(n)a - d_1(n)a^2 - d_2(n)a^3 - \ldots, \quad (2.93)$$

where we have written the operator matrix element as simply $\langle\mathcal{O}(M)\rangle$. The coupling relevant to the operator matrix element is governed by the following beta-function equation,

$$M\frac{\partial a}{\partial M} = \beta(a) = -ba^2(1 + ca + c_2a^2 + c_3a^3\cdots). \quad (2.94)$$

The coefficients $d_i(n)$ depend on the factorization procedure just as the coefficients c_i depend on the renormalization scheme one adopts. Only the subscript-free coefficients i.e. b, c, and d, are FRS invariant.

The coefficient function, $\mathcal{C}_n$, is both RS and FS dependent. It can be written as an expansion of powers of the coupling evaluated at $M = \mu$,

$$\mathcal{C}(Q, M, \mu, \tilde{a}) = 1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3 + \cdots, \quad (2.95)$$

where $\tilde{a} = a(M = \mu)$. One-loop, two-loop and three-loop expressions for the coefficient functions can be found in Refs. [104], [105–113] and [75–77] respectively.

We will return to these equations in chapter 6 when we come to study F_3 moments in more detail. For now we simply state the LO expression for the moments, obtained by taking the LO forms of Eqs. (2.93), (2.94) and (2.95).

$$\mathcal{M}_3(n; Q^2) = A_n\, a^{d(n)/b}, \quad (2.96)$$

where A_n is a set of undetermined FRS invariant non-perturbative parameters. In chapter 6 we extend these expressions to NNLO and then subsequently adapt them to the CORGI and EC representations.

Moments of F_2 in electron-proton scattering

In the case of F_2 moments, we must consider contributions from both singlet and non-singlet operators,

$$\mathcal{M}_2(n;Q^2) = \mathcal{M}_2^{\rm NS}(n;Q^2) + \mathcal{M}_2^{\rm S}(n;Q^2). \tag{2.97}$$

The NS part has a structure identical to that of the F_3 moments. However, the singlet component is more complicated.

The singlet operators of lowest twist mix under renormalization, and hence $\mathcal{M}^{\rm S}(n;Q^2)$ receives contributions from both the quark and gluon operators of Eqs. (2.39) and (2.40). In this case, the moments can be factorized in the form,

$$\mathcal{M}_2^{\rm S}(n;Q^2) = \boldsymbol{C}_n^{\rm S}(Q,a(\mu),\mu,M)\langle\boldsymbol{\mathcal{O}}_n^{\rm S}(M)\rangle. \tag{2.98}$$

$\boldsymbol{C}_n^{\rm S}$ has a quark and a gluon component and the anomalous dimension, $\boldsymbol{\gamma}$, is a 2×2 matrix;

$$\boldsymbol{\gamma}_{n,\rm S}(a) = \begin{pmatrix} \gamma^{qq}(a) & \gamma^{qg}(a) \\ \gamma^{gq}(a) & \gamma^{gg}(a) \end{pmatrix}, \tag{2.99}$$

$$\boldsymbol{C}_n^{\rm S} = (C^q, C^g). \tag{2.100}$$

Consequently, the operator matrix element also has q and g components. The matrix/vector nature of $\boldsymbol{C},\boldsymbol{\mathcal{O}},\boldsymbol{\gamma}$ etc. is denoted by bold typeface. $\boldsymbol{C}$ and $\boldsymbol{\gamma}$ can both be written as expansions in powers of the coupling:

$$\boldsymbol{\gamma}_{n,\rm S}(a) = -\boldsymbol{d}(n)a - \boldsymbol{d}_1(n)a^2 - \boldsymbol{d}_2(n)a^3 - \cdots, \tag{2.101}$$

$$C^q = 1 + r_1^q\tilde{a} + r_2^q\tilde{a}^2 + \cdots, \tag{2.102}$$

$$C^g = r_1^g\tilde{a} + r_2^g\tilde{a}^2 + \cdots. \tag{2.103}$$

The FS dependence of the operator matrix element is described by the following equation,

$$M\frac{\partial\langle\boldsymbol{\mathcal{O}}_n^{\rm S}(M)\rangle}{\partial M} = \boldsymbol{\gamma}_{n,\rm S}(a)\langle\boldsymbol{\mathcal{O}}_n^{\rm S}(M)\rangle. \tag{2.104}$$

These expressions are somewhat harder to solve than their F_3 counterparts; we will

return to them in chapter 7 when we come to study F_2 in more detail. For now, we state the LO result for the moments. In order to obtain this it is necessary to diagonalize $\boldsymbol{d}(n)$; Eqs. (2.98) - (2.103) then yield,

$$\mathcal{M}^{\rm S}(n;Q^2) \;=\; A_n^+ a^{\lambda_+(n)/b} + A_n^- a^{\lambda_-(n)/b}, \tag{2.105}$$

where $\lambda_\pm(n)$ are the eigenvalues of $\boldsymbol{d}(n)$, and $A_n^\pm$ are sets of non-perturbative FRS invariant constants, analogous to A_n.

2.8 Summary

In this chapter, we presented a brief introduction to deep inelastic lepton-hadron scattering. We introduced the naïve parton model and showed how it can give us a qualitative understanding of lepton-hadron interactions in terms of quark distribution functions. We then described how the OPE formalism can be used to calculate QCD corrections to the parton model.

We showed how the RGE for composite operators is derived and demonstrated how it can be used to determine the scale dependence of the moments of structure functions. We then presented a brief discussion of factorization and drew analogies between *it* and the renormalization process. Finally we defined the DIS observables which we will study in this thesis.

Chapter 3

The large-order behaviour of perturbative QCD

In this chapter we discuss the behaviour of perturbation theory at large-orders of the coupling. The validity of perturbation theory depends very much on the convergence properties of the perturbative expansion and up to this point, it has been a tacit assumption that this expansion is absolutely convergent, for sufficiently small values of the coupling. In this chapter, we challenge this assumption. In so doing, we will discover that the perturbative expansions of all QCD observables are *not* convergent, rather they are asymptotic. As a consequence of this, observables can only be unambiguously defined when perturbation theory is supplemented with information from non-perturbative QCD. In this way a bridge is formed between the perturbative and non-perturbative physics.

3.1 Introduction

In 1952, Dyson put forward an intuitive argument suggesting that the perturbative expansions for all QED observables are divergent [114], with the nth-order perturbative coefficients behaving as $n!$. Subsequently, this was confirmed and it was discovered that QFT in general suffers from the same problem. Two sources of this divergence have since been identified: the proliferation of Feynman diagrams at large-orders, and the behaviour, under renormalization, of a single class of diagrams. The latter is referred to as renormalon divergence – it is the most troublesome and arguably the most

interesting of the two [115, 116].

The existence of renormalons has inspired a series of all-orders calculations of the perturbative coefficients for certain observables. Their generic form has been determined, and in some cases their exact form has been obtained in the large-N_f limit. Remarkably, these results allow the full form of the coefficients to be approximated at all-orders, via the leading-b approximation.

These results also confirm the factorial nature of perturbative coefficients. As a consequence of this, we are forced to weaken the assumption that observables can be unambiguously defined by their perturbative expansions. Indeed, by studying the nature of asymptotic series, we are led to conclude that perturbation theory cannot be a complete description of nature, that it can only determine the value of an observable to within a finite maximum accuracy, and that it must be supplemented by non-perturbative physics in order to be completely well-defined.

Once we have understood this, we can use techniques such as Borel resummation to make sense of the factorially divergent sum, and also to extract information on non-perturbative effects. In this way, the study of large order perturbation theory can be used to improve upon the accuracy of standard fixed-order (FO) predictions and also highlights the caveats which apply to the use of FO results. Perhaps most usefully, this also gives us insight into the interplay between perturbative and non-perturbative physics.

The material in this chapter is organized as follows:

We begin by outlining Dyson's argument and how it hints at a factorial form for the perturbative coefficients at large-orders. We then briefly discuss the two sources of divergence so far discovered in QCD: instantons and renormalons. Next, we describe some general properties of asymptotic series and discuss how the Borel transform technology can be used to define such series in a more meaningful way.

We present a discussion of renormalons in QED and QCD, emphasizing how their existence facilitates the obtaining of all-orders leading-N_f results. We then introduce the leading-b approximation and describe how it can be used to approximate the full

form of perturbative coefficients at all-orders. Finally, we describe how the inherent ambiguities associated with perturbative QCD (generated by renormalons) provide a link between perturbative and non-perturbative QCD.

3.2 Dyson's Argument

The potentially divergent behaviour of perturbative coefficients was first highlighted by Dyson. A simple, intuitive argument was presented which suggested that all QED perturbative expansions are divergent [114]. The argument is as follows:

Perturbative corrections to a generic QED observable, $\mathcal{F}(\alpha)$ can always be expressed in the following form,

$$\mathcal{F}(\alpha) = \sum_{n=0}^{\infty} f_n \alpha^{n+1}, \tag{3.1}$$

where $\alpha = e^2/4\pi$ is the QED coupling. If we assume that this series is absolutely convergent for some small *positive* value of α, then the series must have a finite radius of convergence in the complex-α plane. Consequently, the series will also converge for small *negative* values of α.

Consider the physical implications of adopting a negative value for α. The most dramatic consequence is a reversal in the nature of electromagnetic interactions; like charges will attract each other and unlike charges will repel. Let us now examine the effect this would have on a system of N electrons confined to a 'small' region of space. Due to the nature of the potential between two electrons, the energy of this system can always be lowered by simply adding another electron. Moreover, if the region of confinement is small enough, then this drop in the total energy of the system *can* be larger in magnitude than that required to create an e^+e^- pair.

We can now imagine a scenario in which (due to quantum fluctuations) a number of e^+e^- pairs are created from the vacuum and subsequently separate from their partners in such a way that separate concentrations of electrons and of positrons are formed. This configuration of the system reduces its total energy, and this reduction can be such that it cancels out the energy 'borrowed' from the vacuum in order to initially

create the e^+e^- pairs.

From this we conclude that the vacuum is unstable i.e. it can always decay into a state with lower energy. This state is pathological in the sense that eventually an infinite number of pairs would have to be created in order to minimize the energy of the system. Hence it is unimaginable that observables in such a system could be described by convergent perturbative expansions.

If we now reverse our initial line of reasoning, we see that $\mathcal{F}(\alpha)$ cannot converge for negative values of α and hence it must have a zero radius of convergence. This in turn, suggests that $\mathcal{F}(\alpha)$ has a singularity at $\alpha = 0$ and that the series in Eq. (3.1) is at best asymptotic and at worst completely divergent.

The nature of the coefficients

We can shed light on the implications of Dyson's argument for the exact nature of the perturbative coefficients by revisiting the system of N electrons. The total energy of such a system is given by,

$$E \sim NT + \frac{1}{2}N^2V\alpha, \tag{3.2}$$

where T is the mean kinetic energy per electron, V is the mean potential energy between two electrons, and $N^2 \sim \frac{1}{2}N(N-1)$ counts the number of interacting electron pairs. For the physical ($\alpha > 0$) case, there is a stable minimum at $N = 0$. However, for the $\alpha < 0$ case, $N = 0$ is only a local minimum. There is then a peak at $N = N_{\text{crit}}$, where,

$$N_{\text{crit}} \sim \frac{1}{|\alpha|}, \tag{3.3}$$

and above N_{crit}, the energy decreases as $E \sim -N^2$. The existence of no stable minimum is a manifestation of the pathology described above.

From Eq. (3.3), we can infer that the system becomes unstable when N_{crit} electrons are present and hence that Dyson's vacuum instability occurs via the creation of $\sim N_{\text{crit}}$ e^+e^- pairs from the vacuum. A Feynman diagram describing such an event has a factor of $\alpha^{N_{\text{crit}}}$ associated with it, and therefore it is natural to assume that the divergence of $\mathcal{F}(\alpha)$ is related to the behaviour of the set of coefficients, $f_{n \geq N_{\text{crit}}}$. We know that

the series in Eq. (3.1) must experience an initial period of convergence (QED and QCD perturbation series appear to converge at low orders of α). However, at some point this must be followed by a growth in successive terms in the series. Hence, we interpret N_{crit} as a point intermediate between these regions of convergence and divergence. Therefore, at $n \sim N_{\text{crit}}$, we have,

$$f_n \alpha^{n+1} \sim f_{n+1} \alpha^{n+2}, \tag{3.4}$$

$$\Rightarrow \frac{f_{n+1}}{f_n} \sim \frac{1}{\alpha}, \tag{3.5}$$

and then from Eq. (3.3),

$$\frac{f_{n+1}}{f_n} \sim N_{\text{crit}} \sim n. \tag{3.6}$$

This directly implies that $f_n \sim n!$ for $n \gtrsim N_{\text{crit}}$, and we conclude that the divergent nature of $\mathcal{F}(\alpha)$ highlighted by Dyson manifests itself as a factorial growth of the coefficients. Furthermore, we recognize factorial behaviour as signalling that the series is *asymptotic*.

The above argument is not completely rigorous. For instance, it is possible that $\mathcal{F}(\alpha)$ contains non-perturbative terms of the form $e^{-1/\alpha}$ and that these are responsible for the non-analyticity at $\alpha = 0$. However, it is now accepted that QED perturbative series are indeed asymptotic. Furthermore, this problem is also endemic in QCD, with all Green's functions (and therefore all observables) being described by divergent perturbative expansions. In order to be able to make sense of these divergent predictions,[1] we need to know more about asymptotic series. However, before we study this in detail, we first briefly discuss the known mechanisms by which factorial coefficients are generated in QFT.

[1] Care must be taken not to confuse this type of divergence with that which was dealt with using the renormalization procedure. Here, we use the word divergent to refer to the $n!$ behaviour of the coefficients which is present even after the Green's functions have been renormalized. Indeed, in the case of renormalons, it is actually the renormalization process itself which generates the factorial form of the coefficients.

3.3 Sources of divergence

The exact cause of the divergence discovered by Dyson has, so far, remained unclear. By uncovering the mechanisms by which perturbation series become divergent, it is hoped that we can obtain information about the higher-order coefficients and hence, if not solve the problem, then at least quantify it.

Two independent sources of divergence in QFT have so far been identified. The first is the proliferation of Feynman diagrams at large-orders, and this source is intimately linked with the existence of instantons within QFT. The second source is due to a single class of renormalized diagrams, whose contribution grows factorially at large n.

3.3.1 Instanton based divergence

This source has been studied extensively in the case of the generalized quantum anharmonic oscillator [117, 118]. To give an impression of the mechanism by which this type of divergence is generated, we consider the partition function for ϕ^4 theory in $1+0$ dimensions in quantum mechanics,

$$Z(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\phi \, \mathrm{e}^{-\phi^2 - g\phi^4}. \tag{3.7}$$

A cursory examination of $Z(g)$ leads us to conclude that it has an essential singularity at $g = 0$. This is most easily seen by examining the behaviour as $g \to -g$. The 'potential' term in the Lagrangian changes from one with stable minima to one which is not bounded from below.

$Z(g)$ can be expressed as the following perturbative expansion,

$$Z(g) = \sum_{n=0}^{\infty} Z_n g^n, \tag{3.8}$$

and the coefficients Z_n count the number of possible vacuum diagrams contributing at order g^n, in ϕ^4-theory. By expanding the second exponential in Eq. (3.7) and using the definition of the Gamma function given in Eq. (B.1), we can obtain,

$$Z(g) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} d\phi \, \mathrm{e}^{-\phi^2} \sum_{n=0}^{\infty} \frac{(-\phi^4)^n}{n!} g^n$$

$$= \sum_{n=0}^{\infty} \left(\frac{(-1)^n \, \Gamma\left(2n + \frac{1}{2}\right)}{\sqrt{8\pi} \; \Gamma\left(n+1\right)} \right) g^n, \tag{3.9}$$

$$\Longrightarrow \qquad Z_n \;\sim\; \frac{(-4)^n}{4\pi}(n-1)!, \tag{3.10}$$

where in the last step we have taken the large n limit. Barring any significant cancellation between diagrams, and under the reasonable assumption that each diagram contributes with equal magnitude[2] as n grows, then this result identifies the proliferation of Feynman diagrams as a source of divergence in perturbation theory.

A more rigorous analysis of this example [119], confirms this and yields the following result for the general form of Z_n,

$$Z_n \;\simeq\; A \, n^{\gamma} z_i^{-n} n! \left[1 + \mathcal{O}\left(\frac{1}{n}\right) \right]. \tag{3.11}$$

Note that there is an apparent paradox in the results above. The divergence of the perturbative expansion in Eq. (3.8) would seem to be in direct contradiction with the fact that $Z(g)$ is otherwise a perfectly well-defined integral for $g > 0$, and has a finite limit as $g \to 0^+$. The resolution to this paradox lies in the realization that the route taken to obtain Eq. (3.10) is flawed. We have adopted a rather cavalier attitude towards the exchange of the integral and summation symbols in Eq. (3.9). In general, they do not commute and this destroys the convergence properties of what is otherwise a well-defined integral [21,22]. As a result, we can only ever approximate $Z(g)$ asymptotically in the $g \to 0$ limit.

In gauge field theory, cancellations between diagrams will reduce the divergent growth in the coefficients. However, the $n!$ behaviour is still present. Indeed, the proliferation of diagrams at large n is intimately linked with the existence of solutions to the classical equations of motion of a field theory, viz instantons. Specifically, this is caused by the presence of instanton-antiinstanton pairs. This source of divergence is present in QCD (see Refs. [120, 121] and [115, 122]).

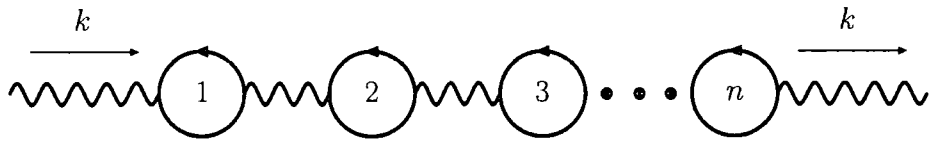


Figure 3.1: A renormalon chain comprising of a string of n fermion loops.

3.3.2 Renormalon based divergence

Renormalons[3] were first identified by 't Hooft as a source of divergence in perturbation theory [40]. They are particular to diagrams containing a string of fermion loops as an internal propagator. These strings are referred to as renormalon chains, and one is shown in Fig. 3.1. The renormalized form of the fermion loops is such that when n of them are coupled together and the momentum running through them is integrated over, the result contains a gamma function which then generates the $n!$ behaviour of the coefficients.

Using Eqs. (1.66) and (1.67) we can evaluate the contributions from such diagrams. For a chain of n fermion loops with momentum k running through it, each bubble will contribute $\sim aN_f k^2 \ln k^2$, and each of the propagators linking the bubbles will contribute k^{-2}. Consequently, the chain will have the form,

$$\text{chain} \approx a^n N_f^n \frac{(\ln k^2)^n}{k^2}. \tag{3.12}$$

It is when we insert a renormalon chain as an internal line in some larger diagram that a factorial form of the coefficients is generated. As an example of this, Fig. 3.2 shows a diagram depicting the renormalon chain correction to the quark-gluon vertex diagram. The diagram in Fig. 3.2 can be understood most easily in terms of the *skeleton expansion*. We refer to the bare quark-gluon vertex diagram as the skeleton diagram and can build the final result from *its* contribution, coupled to the contribution from the renormalon chain.

Crucially, the contribution from the skeleton diagram can be written as a function of k^2/q^2, where q^2 is the external momentum; we denote this function by $\Phi(k^2/q^2)$.

[2]By this we mean the magnitude of the diagram when stripped of factors of the coupling g.

[3]The term renormalon was first coined in the now famous 1977 't Hooft article [40]. Until then the only known sources of divergence were instanton based. The newly discovered source of divergence was known to be rooted in the renormalization process and hence termed renormalon divergence [115].

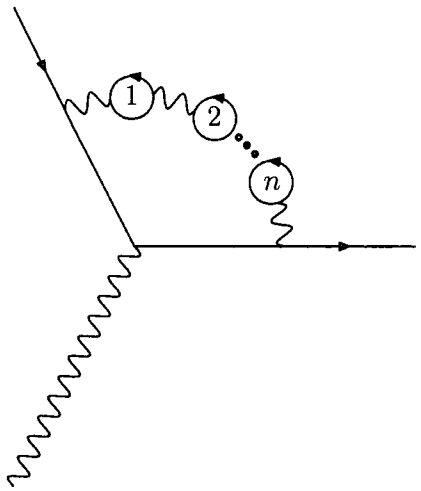


Figure 3.2: A diagram of the type used by 't Hooft to demonstrate the existence of renormalon divergence in perturbation theory.

Furthermore, this function can be written as an expansion in positive powers of k^2/q^2 for $k^2 < q^2$ and as an expansion in negative powers of k^2 for $k^2 > q^2$, with continuity at $k^2 = q^2$.

$$\Phi\left(\frac{k^2}{q^2}\right) \sim \sum_{j>0} \left(\frac{k^2}{q^2}\right)^j \qquad (k^2 < q^2) \tag{3.13}$$

$$\Phi\left(\frac{k^2}{q^2}\right) \sim \sum_{j>0} \left(\frac{q^2}{k^2}\right)^j \qquad (k^2 > q^2) \tag{3.14}$$

Thus, the order a^n contribution to the diagram in Fig. 3.2 will have the form,

$$\text{'t Hooft Diagram} \sim a^n N_f^n \int_0^\infty dk^2 \Phi(k^2/q^2) \left(\ln k^2\right)^n . \tag{3.15}$$

Isolating a single term in the expansions in Eqs. (3.13) and (3.14) in each of the two regions yields,

$$\text{'t Hooft Diagram} \sim a^n N_f^n \left(\int_0^1 \frac{dk^2}{(k^2)^{1-j_1}} \left(\ln k^2\right)^n + \int_1^\infty \frac{dk^2}{(k^2)^{1+j_2}} \left(\ln k^2\right)^n \right) \tag{3.16}$$

Here j_1 and j_2 are positive integers, each of which represents a single term in the expansion of $\Phi(k^2/q^2)$ in each of the regions (Eqs. (3.13) and (3.14)). We have omitted

the q^2 dependence in the above equation[4] and the full result will be a sum over all possible values of j_1, j_2. However, the calculation we have sketched above demonstrates how renormalon chains generate factorially divergent coefficients. Via a change of variables, the terms on the RHS of Eq. (3.16) can be transformed into Gamma functions and hence,

$$\text{'t Hooft Diagram} \sim a^n N_f^n \left((-1)^n j_1^{-n-1} + j_2^{-n-1}\right) n! \tag{3.17}$$

This clearly demonstrates that renormalon chains will generate a factorial contribution to the coefficients of any observable derived from the diagram in Fig. 3.2. Note also, that terms in the expansion for $\Phi(k^2/q^2)$ with the lowest values of j_i will dominate Eq. (3.17).

The (divergent) form of this diagram's contribution has its origin in the renormalization of the initial fermion bubble, specifically its $\ln k^2$ form. Furthermore, this type of divergence is present in all Green's functions of all renormalizable field theories. The discussion above is not completely rigorous and we will discuss renormalons in more detail in a subsequent section.[5]

Clearly the discovery of these two sources of divergence is troubling. However, in order to understand the implications for perturbative predictions, we must first understand more about divergent and asymptotic series.

3.4 Asymptotic series

The revelation that QCD perturbation series are divergent forces us to reassess the relationship between an observable and its perturbative representation. Consider the perturbative expansion of a generic QCD observable, $\mathcal{F}(a)$, written in terms of the coupling $a \equiv \alpha_s/\pi$,

$$\mathcal{F}(a) = \sum_{n=0}^{\infty} f_n a^n. \tag{3.18}$$

[4]The explicit q^2-dependence disappears in the result of any observable calculated from Eq. (3.16). Only q^2-dependence in the form of a remains.

[5]For a detailed review of renormalons, see Ref. [115].

Previously, when convergence of this series was assumed, the only limit on its accuracy lay in our ability to calculate higher-order terms, and crucially the series had a well-defined $n \to \infty$ limit. This is in contrast to the reality of divergent QCD perturbation series. If we assume a typically divergent form of the coefficients ($f_n = n!$) then calculating Eq. (3.18) to some large maximum n would yield a meaningless result with no correspondence to the measured value of the observable. In light of this, it is unclear how we should interpret the equals sign in Eq. (3.18).

Indeed, we can only ever say that the perturbative expansion is *asymptotic* to the exact observable. To understand exactly what we mean by this, we first state the formal definition of an asymptotic series:

Definition

Consider a function $\mathcal{F}(a)$ which is analytic in some domain $\mathcal{D}$, defined by,

$$\mathcal{D} : |\arg(a)| \leq \frac{\theta}{2}, \qquad |a| \leq \rho, \tag{3.19}$$

where $0 < \theta \leq \pi/4$. In perturbation theory, we approximate such functions using an expansion about the point $a = 0$ (Eq. (3.18)). We can then define a function known as the remainder, which evaluates how close the perturbative approximation is to the exact function:

$$R_N = \left| \mathcal{F}(a) - \sum_{n=0}^{N} f_n a^n \right|. \tag{3.20}$$

The series of Eq. (3.18) is then said to be asymptotic to $\mathcal{F}(a)$ if it satisfies the following condition,

$$R_N \leq \left| f_{N+1} a^{N+1} \right|, \tag{3.21}$$

and also diverges for $a \neq 0$. In other words, the error associated with an order a^N approximation must grow more slowly than the next term in the series. In contrast to convergent series, the remainder for an asymptotic series does not vanish as $n \to \infty$, and we can interpret this as the placing of a limit on the accuracy with which an observable can be approximated within perturbation theory.

We know that asymptotic series are characterized by an initial period of convergence,

during which the accuracy of the approximation appears to improve as we move to higher-orders. We also know that for large n, this behaviour is reversed and the sum begins to diverge. We can therefore infer that, intermediate to these two regions, there is an optimal value of n (which we denote by n_{opt}), where the magnitude of successive terms stabilizes with respect to n. From the nature of the initial convergence and from Eq. (3.21), we can infer that the sum truncated at this point represents the best approximation we can achieve.

Fortunately, Eq. (3.21) allows us to estimate the value of n_{opt}. If we assume the following (very general) form of the coefficients,

$$f_n \;=\; A\, n^{\gamma} \left(\frac{1}{z_i}\right) n!, \tag{3.22}$$

then for large n, this can be approximated using Stirling's formula,

$$f_n a^n \;\simeq\; A\sqrt{2\pi} \exp\left\{ n \ln \frac{a}{z_i \mathrm{e}} + n \ln n + \left(\gamma + \frac{1}{2}\right) \ln n \right\}. \tag{3.23}$$

Minimizing the remainder with respect to n will yield the value of n which corresponds to the best approximation,

$$\left.\frac{\partial}{\partial n} \left(f_n a^n\right)\right|_{n=n_{\text{opt}}} \;=\; 0. \tag{3.24}$$

Using this with Eq. (3.23), it can be determined that [123],

$$n_{\text{opt}} \;\simeq\; \frac{-\left(\gamma + \frac{1}{2}\right)}{W_0\left[-\left|\frac{a}{z_i}\right| \left(\gamma + \frac{1}{2}\right)\right]}. \tag{3.25}$$

Here, W_0 is the *principal* branch of the Lambert W function (cf. Eq. (1.116)), and this expression is valid providing $|a/z_i| \left(\gamma + \frac{1}{2}\right) \leq \frac{1}{\mathrm{e}}$. Equation (3.25) can be further simplified by taking its large n limit,

$$n_{\text{opt}} \;\sim\; \left|\frac{z_i}{a}\right|. \tag{3.26}$$

Therefore truncating Eq. (3.18) after $|z_i/a|$ terms will yield the optimum approximation to the full function.

The error associated with truncating the series at $n = N$ is simply equal to the remainder. Hence, from Eq. (3.21), we can determine that the minimum possible error (which

we denote by $\epsilon(a))$ is just $f_n a^n$ evaluated at $n = n_{\rm opt}$. Hence,

$$\begin{aligned} \epsilon(a) &\equiv \left|\min_n (f_n a^n)\right| \\ &= \left|f_{n_{\rm opt}} a^{n_{\rm opt}}\right|. \end{aligned} \tag{3.27}$$

From Eq. (3.23) this can be estimated in the large n limit,

$$\begin{aligned} \epsilon(a) &\simeq |A|\sqrt{2\pi}\frac{{\rm e}^{-n_{\rm opt}-\gamma-\frac{1}{2}}}{(|a/z_i|)^{\gamma+\frac{1}{2}}} \\ &\sim {\rm e}^{-|\frac{z_i}{a}|}. \end{aligned} \tag{3.28}$$

The significance of the values of $n_{\rm opt}$ and $\epsilon(a)$ depends principally on the value of the coupling a. Assuming that $z_i \sim 1$, then in QED we have[6] $n_{\rm opt} \sim 128$ and $\epsilon(a) \sim 10^{-56}$ [32]. Consequently, the predictive power of perturbative QED is not seriously challenged by the large-order behaviour of the perturbative coefficients.

This is not the case in QCD, where we have[7] $n_{\rm opt} \sim 3.5 - 8.5$ and $\epsilon(a) \sim 10^{-2} - 10^{-4}$ [32]. Clearly this poses a problem, not only does it seriously limit the accuracy of perturbation theory (even at NNLO), but it also leads us to question the self-consistency of perturbation series with such poor convergence properties.

Uniqueness

We have seen that as a direct implication of Eq. (3.21), there is an intrinsic error associated with any asymptotic series. Furthermore, another consequence of Eq. (3.21) is that the series in Eq. (3.18) is not asymptotic to the function $\mathcal{F}(a)$ *only*, but is also asymptotic to an infinite class of functions related to $\mathcal{F}(a)$. This can be shown by noting that, for a generic factorial form of the coefficients ($f_n = A z_i^{-n} n!$) the function,

$$\mathcal{F}'(a) = \mathcal{F}(a) + A \exp\left\{-\frac{B}{a^k}\right\}, \tag{3.29}$$

[6] We take the coupling to be $\alpha_{QED}(Q = m_Z)$.
[7] Here we take α_{QCD} at $Q = \sqrt{3}GeV$ and $Q = m_Z$.

also satisfies the bound in Eq. (3.21), provided that $B\cos(\theta/4) > z_i$, $A \in \mathbb{R}$, $B > 0$ and $0 < k\theta < \pi$ [124]. Therefore, the series in Eq. (3.18) is asymptotic to a class of functions $\mathcal{F}'(a)$ defined by the above equation, and from this we conclude that the series in Eq. (3.18) does not define a unique function.[8] This implies that there is some additional information about the observable which is inaccessible to perturbation theory.

In light of this, we rewrite Eq. (3.18) as,

$$\mathcal{F}(a) \approx \sum_{n=0}^{\infty} f_n a^n, \tag{3.30}$$

where the '$\approx$' symbol means 'is asymptotic to'. The important conclusion that we draw from this is that the asymptotic nature of perturbation theory means that *it* cannot alone define observables in an unambiguous (and hence physically consistent) manner. The presence of such ambiguity is troubling. However, we note that the ambiguity (the additional term on the RHS of Eq. (3.29)) is non-perturbative in structure. We shall see later how, when non-perturbative effects are taken into account, we can recover self-consistent predictions for observables within QCD.

3.4.1 The Borel transform

The existence of an initial period of convergence for asymptotic series implies that the subset of coefficients, $f_{n<n_{\text{opt}}}$, must hold some information about the exact function. What though of the terms beyond this period of convergence, $f_{n>n_{\text{opt}}}$? Although they have a divergent form, they must still hold information about the exact function, since they are obtained directly from it. Indeed, the large-order behaviour of the perturbative coefficients (i.e. their behaviour beyond the period of convergence) can enlighten us as to aspects of the perturbative and non-perturbative sectors of QCD.

Yet, how can information about the exact function be extracted from the large n coefficients when the series effectively satisfies $\sum_{n=0}^{\infty} f_n a^n = \infty$? It is obvious that studying the series in this form is not practical. We can, however, endow the series with some meaning via the Borel transform.

[8] In contrast, the reverse is not true. Each function $\mathcal{F}(a)$ defines a unique asymptotic series.

We define the Borel transform of the series in Eq. (3.18) as:

$$B[\mathcal{F}](z) = \sum_{n=0}^{\infty} \frac{f_n}{n!} z^n. \tag{3.31}$$

We observe that this new series has improved convergence properties; any factorial divergence is cancelled out. Indeed, if we compare the radius of convergence of the Borel transform, ρ_B, with that of the original series, ρ, using the Cauchy root test:

$$\rho_B^{-1} = \lim_{n\to\infty} \sqrt[n]{|f_n|/n!}, \qquad \rho^{-1} = \lim_{n\to\infty} \sqrt[n]{|f_n|}, \tag{3.32}$$

we see that a series with $\rho = 0$ becomes a series with $\rho_B =$ finite, and a series with $\rho =$ finite becomes a series with $\rho_B = \infty$. In effect, the Borel transform expands the radius of convergence such that a factorially divergent perturbation series will have a Borel transform with a finite radius of convergence. Redefining the series in this way allows us to study the $n \to \infty$ behaviour in a more meaningful way.

The original series $\mathcal{F}(a)$ can be recovered from the Borel transform by means of a special Laplace transform, as we shall now demonstrate. Inserting a factor of $1 = \Gamma(n+1)/n!$ into Eq. (3.18) gives,

$$\mathcal{F}(a) = \sum_{n=0}^{\infty} f_n a^n \frac{\Gamma(n+1)}{n!} = \sum_{n=0}^{\infty} \frac{f_n a^n}{n!} \int_0^{\infty} e^{-z} z^n dz. \tag{3.33}$$

Moving the sum and integral symbols through each other, and making a change of variables, then gives the *Borel representation* of $\mathcal{F}(a)$,

$$\begin{aligned} \tilde{\mathcal{F}}(a) &= \int_0^{\infty} e^{-z} dz \sum_{n=0}^{\infty} \frac{f_n}{n!} (az)^n \\ &= \frac{1}{a} \int_0^{\infty} e^{-z/a} B[\mathcal{F}](z) dz. \end{aligned} \tag{3.34}$$

This defines the Borel sum of $\mathcal{F}(a)$. It is tempting to think that this expression is equal to that given in Eq. (3.18) - allowing us to omit the tilde symbol. However, the switching of the integral and sum signs is not strictly valid in general, and hence the above derivation is not entirely mathematically rigorous. In fact, the two series definitions $\mathcal{F}(a)$ and $\tilde{\mathcal{F}}(a)$ are asymptotic to each other, $\tilde{\mathcal{F}}(a) \approx \mathcal{F}(a)$. In spite of this, in most cases the Borel sum can be used to give more precise meaning to asymptotic series. Through it, we can obtain a well-defined function which is asymptotic to the observable.

Formally, the series in Eq. (3.18) is *Borel summable* if it satisfies the following requirements [124]:

- The Borel transform, $B[\mathcal{F}](z)$, converges inside some radius, $\delta > 0$.
- The Borel sum converges for some $a \neq 0$.
- $B[\mathcal{F}](z)$ can be analytically continued to a strip of the complex z-plane with non-vanishing width, bisected by the positive real semi-axis.

One of the most useful features of the Borel transform is the fact that it can be used to classify different asymptotic series. To illustrate this, consider the following two series,

$$\mathcal{F}_1(a) = \sum_{n=0}^{\infty}(-1)^n n!\, a^n, \qquad \mathcal{F}_2(a) = \sum_{n=0}^{\infty} n!\, a^n. \tag{3.35}$$

Using Eq. (3.31) we can obtain the Borel transforms of these series, and these can be resummed as Taylor expansions, yielding,

$$f_n = (-1)^n n! \qquad f_n = n!, \tag{3.36}$$

$$\downarrow \qquad\qquad \downarrow$$

$$B[\mathcal{F}_1](z) = \frac{1}{1+z} \qquad B[\mathcal{F}_2](z) = \frac{1}{1-z}, \tag{3.37}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\tilde{\mathcal{F}}_1(a) = \frac{1}{a}\int_0^{\infty} e^{-z/a}\frac{1}{1+z}dz \qquad \tilde{\mathcal{F}}_2(a) = \frac{1}{a}\int_0^{\infty} e^{-z/a}\frac{1}{1-z}dz. \tag{3.38}$$

We can see that $B[\mathcal{F}_1](z)$ and $B[\mathcal{F}_2](z)$ have poles on the real axis of the complex z-plane; this is a familiar feature of the Borel transforms of asymptotic series. They are a manifestation of the divergent nature of the series expansion of $\mathcal{F}_i(a)$ (or alternatively, of its zero radius of convergence). The position and nature of the singularities are indicative of the specific type of divergence. In general, an alternating sign divergence generates poles on the negative real semi-axis, whereas a fixed sign divergence generates poles on the positive real semi-axis. The difference is crucial since it is only singularities on the positive real semi-axis which lie within the range of integration of the Borel integral. These poles cause problems in defining the Borel sum. Indeed, a prescription must be chosen for negotiating them, and this freedom of choice generates an ambiguity in the evaluation of the Borel sum. In general, divergent series generate poles or

branch points on the real axis. A common choice of prescription for negotiating such singularities occurring on the integration contour is the Cauchy principal value (PV) prescription. This prescription is defined by the following expression,

$$\mathrm{PV}\int_a^b f(x)dx \equiv \lim_{\epsilon\to 0^+}\left[\int_a^{c-\epsilon} f(x)dx + \int_{c+\epsilon}^b f(x)dx\right]. \tag{3.39}$$

$\tilde{\mathcal{F}}_1(a)$ and $\tilde{\mathcal{F}}_2(a)$ can be evaluated in terms of the exponential integral function, $\mathrm{Ei}(x)$, defined in Eq. (B.6),

$$\tilde{\mathcal{F}}_1(a) = -\frac{\mathrm{e}^{1/a}}{a}\mathrm{Ei}\left(-\frac{1}{a}\right), \qquad \tilde{\mathcal{F}}_2(a) = \frac{\mathrm{e}^{-1/a}}{a}\mathrm{Ei}\left(\frac{1}{a}\right). \tag{3.40}$$

$\tilde{\mathcal{F}}_1(a)$ requires no prescription and can easily be evaluated. However, the exponential integral in $\tilde{\mathcal{F}}_2(a)$ must be regulated and hence this Ei function refers to the PV regulated exponential integral function (see section B.3).

We conclude from these examples that an alternating sign divergent series can be resummed using the Borel sum technique, and we can ascribe to the series a well-defined and unique value. For fixed sign divergence however the situation is less encouraging. Singularities in the Borel transform make the result ambiguous. This ambiguity corresponds directly to the possible non-uniqueness of asymptotic series discussed in section 3.4 and expressed in Eq. (3.29). However, via the Borel transform, we can resum perturbative expansions to all-orders and hence represent the corresponding observables asymptotically.

Finally, we study the Borel transform for a more general and realistic asymptotic series, in which the coefficients have the form in Eq. (3.22). We have,

$$B[f](z) = A\sum_{n=0}^{\infty} z_i^{-n} n^{\gamma} z^n. \tag{3.41}$$

Using the following result (obtained by using the binomial theorem and then taking the large n limit),

$$(1-z)^{-t} = \sum_{n=0}^{\infty}\frac{\Gamma(n+t)}{\Gamma(n+1)\Gamma(t)}\,z^n,$$

$$\approx \sum_{n=0}^{\infty} \frac{n^{t-1}}{(t-1)!} z^n, \tag{3.42}$$

we can resum Eq. (3.41) to obtain,

$$B[\mathcal{F}](z) \approx A \frac{\Gamma(\gamma+1)}{\left(1-\frac{z}{z_i}\right)^{\gamma+1}}. \tag{3.43}$$

This Borel transform is similar to those of Eq. (3.37). However, the position of singularities on the real axis is now determined by the sign and magnitude of z_i. The nature of the singularities is determined by γ, with non-integer γ corresponding to a branch point.

All QCD Green's functions are related to each other through the Schwinger-Dyson equations [125–127]. Furthermore, the Green's functions are related in such a way that the positions of singularities in the Borel plane are preserved [40], although the nature of the singularities can be altered. Consequently, singularities in the Borel transform of one Green's function propagate throughout the entire theory, infecting all Green's functions. Therefore in general, the Borel transforms of all QCD Green's functions and observables have branch points along the real axis due to renormalons and instantons [128, 129].

3.5 Renormalons

Having identified renormalons as the principle source of divergence in perturbative QCD, and having developed a set of tools (in the form of the Borel transform) to evaluate asymptotic series, we now discuss the origin of renormalons in more detail.

As stated previously, the source of renormalon divergence is the existence of a class of Feynman diagrams, present at all-orders, which when renormalized are found to have a magnitude proportional to $n!$. To fully appreciate the mechanism by which this occurs, we return to the familiar example of the fermion loop contribution to the vacuum polarization function, depicted in Fig. 1.4. From Eqs. (1.66) - (1.67) and (1.83) we have:

$$-i\Pi_{ab}^{\mu\nu}(k^2) \equiv \delta_{ab}(k^\mu k^\nu - g^{\mu\nu}k^2)\Pi_0(k^2), \tag{3.44}$$

with

$$\Pi_0(k^2) \;=\; i\frac{a}{4}\frac{4T_R}{3}\ln\left(\frac{k^2}{Q^2}\right). \tag{3.45}$$

Here, $Q^2 \equiv -q^2 > 0$ is the space-like energy scale, and we work in the V scheme, which is $\overline{\rm MS}$ with $\mu^2 = {\rm e}^{-5/3}Q^2$. In a sense, this is purely a QED (Abelian) calculation: we would have to include gluon and ghost corrections in order to obtain the full one-loop function. However, the fermion bubble *is* a component of the full QCD result and moreover, we will see that it is of special significance.

We denote the chain of n fermion loops depicted[9] in Fig. 3.1 by $B^{\mu\nu}_{(n)}(k^2)$. We can evaluate the form of $B^{\mu\nu}_{(n)}(k^2)$ using Eq. (3.44), together with the form of the gluon propagator given in Eq. (1.50). We find,

$$\begin{aligned} B^{\mu\nu}_{(n)}(k^2) \;&=\; -i(-1)^n(P^{\mu\beta_1})(\Pi_{\beta_1\alpha_2})(P^{\alpha_2\beta_2})(\Pi_{\beta_2\alpha_3})\ldots(P^{\alpha_n\beta_n})(\Pi_{\beta_n\alpha_{n+1}})(P^{\alpha_{n+1}\nu}) \\ &=\; -i(-1)^n\prod_{k=1}^{n}\left[\left(P^{\alpha_k\beta_k}\right)\left(\Pi_{\beta_k\alpha_{k+1}}\right)\right](P^{\alpha_{n+1}\nu}). \end{aligned} \tag{3.46}$$

Here, we have made the identification $\alpha_1 \equiv \mu$, and $-iP^{\mu\nu}(k^2) \equiv \tilde{D}^{\mu\nu}_{ab}(k)$ represents the photon/gluon propagator[10] (see Eq. (1.50)). The above expression can be evaluated using the following results,

$$P^{\alpha_k\beta_k}(k^2)\Pi_{\beta_k\alpha_{k+1}}(k^2) \;=\; \frac{1}{k^2}\Pi^{\alpha_k}{}_{\alpha_{k+1}}(k^2), \tag{3.47}$$

$$\prod_{k=1}^{n}\left[\Pi^{\alpha_k}{}_{\alpha_{k+1}}(k^2)\right] \;=\; \left(\Pi^{\alpha_1}{}_{\alpha_{n+1}}(k^2)\right)\left(\Pi_0^{n-1}(k^2)\right)[k^2]^{n-1}. \tag{3.48}$$

Hence,

$$\begin{aligned} B^{\mu\nu}_{(n)}(k^2) \;&=\; \frac{(-1)^n}{k^2}(\Pi_0)^{n-1}\left(\Pi^{\alpha_1}{}_{\alpha_{n+1}}\right)(-iP^{\alpha_{n+1}\nu}) \\ &=\; (-1)^n(\Pi_0)^n\left(\frac{-i}{k^2}\right)\left[g^{\mu\nu}-\frac{k^\mu k^\nu}{k^2}\right], \end{aligned} \tag{3.49}$$

We can sum this over all values of n to obtain the all-orders renormalon chain contri-

[9] Hereafter, we use the photon line symbol to denote both photons and gluons.
[10] We omit colour indices in the following demonstration.

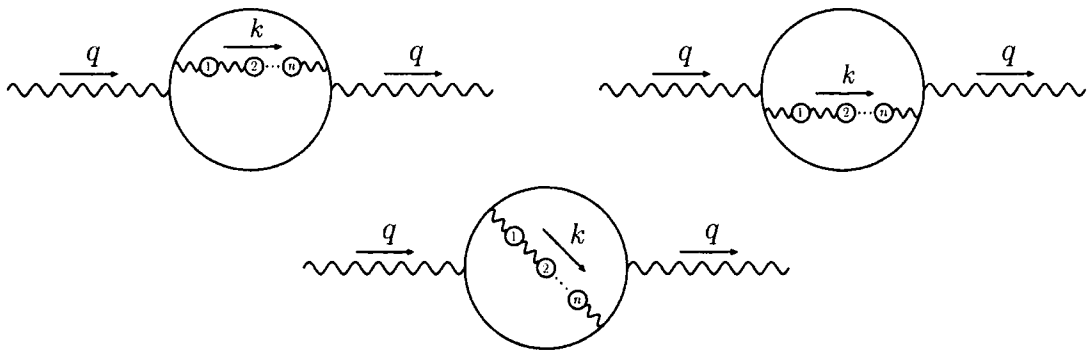


Figure 3.3: One-chain contributions to the vacuum polarization function.

bution to the propagator, which we denote by $B_{\mu\nu}(k^2)$,

$$\begin{aligned} B_{\mu\nu}(k^2) &= \sum_{n=1}^{\infty} B_{\mu\nu}^{(n)}(k^2) \\ &= \left(\frac{-i}{k^2}\right)\left[g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right]\sum_{n=1}^{\infty}(-1)^n\left(\Pi_0(k^2)\right)^n \\ &= \left(\frac{-i}{k^4}\right)\left[k^2 g_{\mu\nu} - k_\mu k_\nu\right]\frac{-\Pi_0(k^2)}{1+\Pi_0(k^2)}. \end{aligned} \tag{3.50}$$

The complete gauge boson propagator is defined as the sum over all possible diagrams with two external photon lines only. Calculating this exactly would correspond to calculating all possible diagrams in perturbation theory – something we wish to avoid. However, of all the diagrams which contribute to the full vacuum polarization function, the above one-chain result has the highest power of N_f at each order of a. In the large-N_f limit, Eq. (3.50) *is* the full vacuum polarization function. Therefore, by inserting $B^{\mu\nu}(k^2)$ as an internal line in a larger skeleton diagram, we can identify the leading-N_f components of that diagram. Moreover, we can use the relatively simple expression above to evaluate that diagram at all-orders of a.

The renormalon contributions to the Adler D-function

Armed with the leading N_f result for the gluon/photon propagator, we now demonstrate how single-chain contributions to QCD observables can be evaluated, using the

Adler D-function (defined in section 1.10.1) as an example.

Corrections to the Adler D-function are calculated from corrections to the vacuum polarization function. The one-chain contributions to such corrections are shown in Fig. 3.3 and these will have the form:

$$\Pi(Q^2) \sim a\sum_{n=0}^{\infty}\int\frac{d^4k}{(2\pi)^4}\frac{d^4p}{(2\pi)^4}\left[B^{\sigma\rho}_{(n)}(k^2)\mathrm{Tr}\left(\gamma_\nu\frac{1}{\not p+\not q+\not k}\gamma_\rho\frac{1}{\not p+\not q}\gamma_\mu\frac{1}{\not p}\gamma_\sigma\frac{1}{\not p+\not k}\right)\right.$$
$$\left. + \; 2B^{\sigma\rho}_{(n)}(k^2)\mathrm{Tr}\left(\gamma_\nu\frac{1}{\not p+\not q}\gamma_\mu\frac{1}{\not p}\gamma_\sigma\frac{1}{\not p+\not k}\gamma_\rho\frac{1}{\not p}\right)\right] \tag{3.51}$$

$$\sim a\sum_{n=0}^{\infty}\int\frac{d^4k}{(2\pi)^4}\frac{d^4p}{(2\pi)^4}\left[B^{\sigma\rho}_{(n)}(k^2)X_{\nu\rho\mu\sigma}+2B^{\sigma\rho}_{(n)}(k^2)\bar{X}_{\nu\mu\sigma\rho}\right]. \tag{3.52}$$

Here, k is the momentum running through the renormalon chain and p is the momentum running around the outer fermion loop. This can be rewritten in such a way that the chain and skeleton diagram contributions are separated,

$$\Pi(Q^2) \sim \underbrace{\left(\int\frac{d^4k}{(2\pi)^4}B_{\sigma\rho}(k^2)\right)}_{\text{renormalon chain}}\underbrace{\left(\int\frac{d^4p}{(2\pi)^4}\left[X^{\nu\rho\mu\sigma}+2\bar{X}^{\nu\mu\sigma\rho}\right]\right)}_{\text{skeleton diagram contribution}}. \tag{3.53}$$

Omitting the tensor structure, and performing the p integration will yield a result of the form,

$$\Pi(Q^2) \sim \int\frac{d^4k}{(2\pi)^4}B_{\sigma\rho}(k^2)F\left(k^2,k\cdot q,q^2\right). \tag{3.54}$$

To obtain the Adler D-function, we take the logarithmic derivative of the above equation with respect to Q^2. The Q^2-dependence of $B_{\sigma\rho}(k^2)$ occurs in the form of factors of k^2/Q^2. Therefore, via a change of variables, we can transfer all of the Q^2-dependence of Eq. (3.54) into the function F, and hence when we differentiate with respect to Q^2, the renormalon chain term is preserved. Consequently, we can write the Adler D-function as,

$$\mathcal{D}(Q^2) = a\sum_{n=0}^{\infty}\int_0^\infty\frac{dk^2}{k^2}\Phi\left(\frac{k^2}{Q^2}\right)\left[a\frac{N_fT_R}{3}\ln\left(\frac{k^2}{Q^2}\right)\right]^n. \tag{3.55}$$

The kernel[11] $\Phi(k^2/Q^2)$ is derived from $F\left(k^2,k\cdot q,q^2\right)$, and represents the contribution

[11] Here, the form of $\Phi(k^2/Q^2)$ is different to that for the example given in section 3.3.2.

from the skeleton diagram [130]. The term in the square brackets comes solely from the renormalon chain and can be obtained directly from Eq. (3.50). $\Phi(k^2/Q^2)$ can be written as an expansion in positive powers of k^2/Q^2 for $k^2 < Q^2$ and in negative powers for $k^2 > Q^2$, with continuity at $k^2 = Q^2$. The leading terms in these expansions[12] are [131, 132],

$$\Phi\left(\frac{k^2}{Q^2}\right) = \frac{k^2}{Q^2}\left(2\frac{k^2}{Q^2}\right) \qquad \text{for } k^2 < Q^2, \text{ (IR)}, \tag{3.56}$$

$$\Phi\left(\frac{k^2}{Q^2}\right) = \frac{k^2}{Q^2}\left[\frac{10}{27} + \frac{4}{9}\ln\left(\frac{k^2}{Q^2}\right)\right]\frac{Q^4}{k^4} \qquad \text{for } k^2 > Q^2, \text{ (UV)}. \tag{3.57}$$

From Eqs. (3.55) - (3.57) we can directly obtain the leading-b contribution to the Borel transform of the Adler D-function. Splitting the range of integration in Eq. (3.55) at $k^2 = Q^2$, and using the changes of variables $k^2 = Q^2\mathrm{e}^{-z/2a}$ and $k^2 = Q^2\mathrm{e}^{z/2a}$ in Eqs. (3.56) and (3.57) respectively yields,

$$\mathcal{D}(Q^2) = \sum_{n=0}^{\infty}\int_0^{\infty} dz\, \mathrm{e}^{-z/a} \times\left[\left(-\frac{N_f T_R z}{6}\right)^n + \frac{10}{27}\left(\frac{N_f T_R z}{3}\right)^n + \frac{4}{9}(n+1)\left(\frac{N_f T_R z}{3}\right)^n\right]. \tag{3.58}$$

Here, we have performed integration by parts on the ln term in Eq. (3.57). From this we can determine the Borel transform,

$$B[\mathcal{D}](z) = \frac{1}{1+\frac{N_f T_R z}{6}} + \frac{10/27}{1-\frac{N_f T_R z}{3}} + \frac{4/9}{\left(1-\frac{N_f T_R z}{3}\right)^2}. \tag{3.59}$$

Equation (3.55) can also be used to obtain the form of the coefficients d_n, defined in Eq. (1.187). Making the change of variables $k^2 = Q^2\mathrm{e}^{\pm t}$, yields,

$$\mathcal{D}(Q^2) = a\sum_{n=0}^{\infty}\left(a\frac{N_f T_R}{3}\right)^n\int_0^{\infty} dt\, t^n\left[2\mathrm{e}^{-2t}(-1)^n + \left(\frac{10}{27}+\frac{4}{9}t\right)\mathrm{e}^{-t}\right] \tag{3.60}$$

$$= a\sum_{n=0}^{\infty}\left(a\frac{N_f T_R}{3}\right)^n\left[\left(\frac{-1}{2}\right)^n + \left(\frac{22+12n}{27}\right)\right]n!, \tag{3.61}$$

[12] Note that only when all terms (or certain special subsets of terms) in this expansion are considered do we have continuity at $k^2 = Q^2$.

and this confirms that renormalons generate factorial contributions to the coefficients,

$$d_n = \left(\frac{T_R N_f}{3}\right)^n \left[\left(\frac{-1}{2}\right)^n + \left(\frac{22+12n}{27}\right)\right] n!. \tag{3.62}$$

There are several points to note about the above results:

- Higher-order terms in Eqs. (3.56) and (3.57) will generate terms in the Borel transform with singularities further away from the origin than those in Eq. (3.59). Therefore, their contributions to the coefficients will be less dominant than that of Eq. (3.61) but will still have a factorial form. The full form of $\Phi(k^2/Q^2)$ is known for several observables and from it we can obtain the full leading-N_f Borel transform.

- Equations (3.59) and (3.61) are valid and gauge-invariant results in QED. However, in QCD, resumming the leading-N_f terms in this way generates a result which does *not* exhibit asymptotic freedom. This problem can be overcome by converting to an expansion in powers of the first coefficient of the beta-function equation, b. We will describe this in more detail in the next section.

3.6 All-orders results for QCD observables

The existence of renormalons in QFT has inspired a number of all-orders calculations of the most dominant (factorially) divergent components of the perturbative coefficients. The results from these are in the form of the leading-N_f component of the perturbative coefficients at all-orders. They are essentially QED results, in that their diagrammatic origin does not include contributions from gluon self-coupling or ghost components. However, such a result also corresponds to a significant subset of the full set of terms present in the equivalent QCD result. Moreover, the significance of these terms can be further enhanced by means of the leading-b approximation.

3.6.1 The leading-N_f expansion

The coefficients of any QCD perturbative expansion can be written as an expansion in powers of the number of fermions, N_f. For example, in the case of the Adler D-function,

$$d_n = d_n^{[n]} N_f^n + d_n^{[n-1]} N_f^{n-1} + \ldots + d_n^{[0]}, \tag{3.63}$$

and similar expressions also hold for the coefficients of the GLS and Bjorken sum rules, k_n and u_n (see section 2.7.1).

The number of fermion loops present in a diagram dictates which coefficient in Eq.(3.63) it contributes to. Furthermore, the all-orders results obtained from single renormalon chain insertions correspond to the maximum number of powers of N_f one can have in a diagram for each power of a. That is, we cannot increase the number of powers of N_f (the number of fermion loops) in the diagrams in Fig. 3.3 without introducing at least one extra power of a; diagrams with two chains inserted will be $\mathcal{O}(1/N_f)$ suppressed. We therefore identify the single-chain diagrams of Fig. 3.3 as the sole contribution to $d_n^{[n]}$.

As a consequence of this, the large-N_f expansion and renormalons are inextricably linked. Full renormalon calculations of the type sketched in section 3.5 will yield exact results for $d_n^{[n]}$ (or equivalently, $u_n^{[n]}$ or $k_n^{[n]}$ etc.) for all n. Such results have been obtained in the cases of the Adler D-function, the sum rules of section 2.7.1, and the $R_{e^+e^-}$ ratio, amongst others.

Progress in large-N_f calculations [133] has led Broadhurst to develop the following elegant method which allows the values of the $d_n^{[n]}$ to be easily extracted at all-orders [134]: We introduce the QED Gell-Mann-Low function which at leading-N_f has the form,

$$\Psi_n^{[n]} = \frac{3^{2-n}}{2} \left(\frac{d}{dx}\right)^{n-2} P(x)\Big|_{x=1}, \tag{3.64}$$

where

$$P(x) = \frac{32}{3(1+x)} \sum_{k=2}^{\infty} \frac{(-1)^k k}{(k^2-x^2)^2}. \tag{3.65}$$

$\Psi_n^{[n]}$ can be explicitly evaluated in closed form [134],

$$\frac{\Psi_n^{[n]}}{(n-2)!} = \frac{(n-1)}{(-3)^{n-1}}\Bigg[-2n+4-\frac{n+4}{2^n} + \frac{16}{n-1}\sum_{n^2>s>0} s(1-2^{-2s})(1-2^{2s-n})\zeta_{2s+1}\Bigg], \tag{3.66}$$

and using this result one can then obtain the leading-N_f result for the QCD Adler D-function. In the $\overline{\rm MS}$ scheme with $\mu^2 = Q^2$ we have [135],

$$d_n^{[n]} = 2T_R^n n! \sum_{m=0}^{n} \frac{(-\frac{5}{9})^m}{m!} \frac{\Psi_{n+2-m}^{[n+2-m]}}{(n-m)!}. \tag{3.67}$$

This result becomes particularly simple if we convert it into the V-scheme (due to the absence of a finite part in the renormalized form of the one-loop vacuum polarization function, cf. Eqs. (1.66) and (1.67)). We find,

$$d_n^{[n]} = 2T_R^n \Psi_{n+2}^{[n+2]}, \tag{3.68}$$

and this gives us access to the leading-N_f components of d_n for all n, allowing us, in principle, to evaluate the sum,

$$\sum_{n=0}^{\infty} d_n^{[n]} N_f^n a^{n+1}. \tag{3.69}$$

This result can be applied to both QED *and* QCD. However, the QCD result can be given added meaning via the leading-b expansion.

3.6.2 The leading-b expansion

All-orders renormalon inspired results can play a more meaningful role in QCD if we convert them into the leading-b expansion form, whereby the coefficients are written as an expansion in powers of the first coefficient of the beta-function equation, b:

$$d_n = d_n^{(n)} b^n + d_n^{(n-1)} b^{n-1} + \ldots + d_0^{(0)}. \tag{3.70}$$

This discussion is different depending on whether we are talking about QED or QCD, due to the difference in their respective beta-function equations:

$$b_{QED} = -\frac{2}{3}N_f, \qquad b \equiv b_{QCD} = \frac{1}{6}(33 - 2N_f). \tag{3.71}$$

Note that $b_{QED} < 0$ and $b_{QCD} > 0$. These expressions can be rearranged to give N_f in terms of b,

$$N_f = -\frac{3}{2}b_{QED}, \qquad N_f = -3b + 33/2, \tag{3.72}$$

and so by substitution into Eq. (3.63), we can determine that the leading-N_f and leading-b coefficients in QCD are related by $d_n^{(n)} = (-3)^n d_n^{[n]}$, with a similar expression for QED. The sub-leading terms in Eq.(3.70) can be easily obtained from Eqs.(3.63) and (3.72).

In QED there is a firm diagrammatical interpretation of $d_n^{(n)}$, since there is a direct correspondence between leading-N_f and leading-b diagrams. However, in QCD, this is not the case as $d_n^{(n)}b^n$ contains not only the leading-N_f diagrams, but also an expansion in colour factors. Consequently, the diagrammatic origin of the leading b terms is less clear. However, using the background-field method [136], or the Pinch [137–141] technique, we can identify the component of the gluon propagator proportional to $b\ln(-k^2/\mu^2)$ and from this we can build the leading-b terms in the coefficients in a gauge invariant manner.

However, the benefits of using the leading-b expansion over its leading-N_f equivalent are twofold: firstly, the leading-b term in Eq. (3.70) tends to dominate over the sub-leading terms, and therefore we can use this term to approximate the full form of d_n. Secondly, the sign of b determines the asymptotic behaviour of a field theory, and consequently, is indicative of the UV and IR behaviour. It is therefore more meaningful to discuss the Borel transform in the leading-b form.

All-orders leading-b results

The leading-N_f Borel transform of the Adler D-function can be obtained from Eq. (3.68). Borel transforms for several other observables have also been obtained including the GLS, polarized and unpolarized Bjorken sum rules and the $R_{e^+e^-}$ ra-

tio. These can easily be converted into leading-b results by making the replacement $N_f \to -3b$ in the case of QCD (and $N_f \to (-3/2)b_{QED}$ in the case of QED). In this form, the Borel transforms represent the summation of leading-b terms in Eq. (3.70), which we denote as $\mathcal{D}^{(L)}(a)$,

$$\mathcal{D}^{(L)}(a) = \sum_{n=0}^{\infty} d_n^{(L)} a^{n+1}, \tag{3.73}$$

where we have denoted the leading-b component of d_n as $d_n^{(L)} \equiv d_n^{(n)} b^n$.

We can convert the results of Eqs. (3.58), (3.59) and (3.62) into a leading-b form. We have,

$$\mathcal{D}(Q^2) = \sum_{n=0}^{\infty} \int_0^{\infty} dz\, \mathrm{e}^{-z/a} \left[\left(\frac{bz}{4}\right)^n + \frac{10}{27}\left(-\frac{bz}{2}\right)^n + \frac{4}{9}(n+1)\left(-\frac{bz}{2}\right)^n \right], \tag{3.74}$$

$$\Rightarrow \quad B[\mathcal{D}](z) = \frac{1}{1-\frac{bz}{4}} + \frac{10/27}{1+\frac{bz}{2}} + \frac{4/9}{\left(1+\frac{bz}{2}\right)^2}. \tag{3.75}$$

and consequently,

$$d_n = b^n \left[\left(\frac{1}{4}\right)^n + \left(-\frac{1}{2}\right)^n \left(\frac{22+12n}{27}\right) \right] n!. \tag{3.76}$$

For QCD we have $b > 0$ and hence we conclude that the UV component of $\Phi(k^2/q^2)$ in Eq. (3.57) gives rise to renormalon singularities on the negative real semi-axis in the Borel plane and correspondingly generates an alternating sign factorially divergent contribution to the coefficients. In contrast, the IR component of $\Phi(k^2/q^2)$ in Eq. (3.56) generates singularities on the positive real semi-axis and fixed sign factorial divergence.

The full form of the leading-b Borel transform of the Adler D-function can be obtained from Eqs. (3.64) - (3.68), or alternatively by considering the full form of the expansion in Eqs. (3.56) - (3.57). We find [142, 143],

$$\begin{aligned} B[\mathcal{D}^{(L)}](z) = & \sum_{n=1}^{\infty} \frac{A_0(n) - A_1(n) z_n}{\left(1+\frac{z}{z_n}\right)^2} + \frac{A_1(n) z_n}{\left(1+\frac{z}{z_n}\right)} \\ & + \sum_{n=1}^{\infty} \frac{B_0(n) + B_1(n) z_n}{\left(1-\frac{z}{z_n}\right)^2} - \frac{B_1(n) z_n}{\left(1-\frac{z}{z_n}\right)}, \end{aligned} \tag{3.77}$$

where

$$A_0(n) = \frac{8}{3}\frac{(-1)^{n+1}(3n^2+6n+2)}{n^2(n+1)^2(n+2)^2}, \qquad A_1(n) = \frac{8}{3}\frac{b(-1)^{n+1}(n+\frac{3}{2})}{n^2(n+1)^2(n+2)^2},$$

$$B_0(1) = 0, \qquad B_1(1) = 0,$$

$$B_0(2) = 1, \qquad B_1(2) = -\frac{b}{4},$$

$$B_0(n) = -A_0(-n), \qquad B_1(n) = -A_1(-n), \qquad \text{for } n \geq 3. \tag{3.78}$$

and where $z_n = 2n/b$. These definitions coincide with those of Ref. [143], except for $B_1(2) = -\frac{b}{4}$. The purpose of the slight change of definition is to make more explicit the single and double pole structure. We can also obtain leading-b Borel transforms for the GLS and polarized Bjorken sum rules. These are much simpler, with only a finite number of simple poles and no double poles. We state these forms in chapter 4 when we come to study these all-orders results in more detail.

Using Eq. (1.197), we can also obtain the Borel transform for $R_{e^+e^-}$ [142, 143],

$$B[\mathcal{R}^{(L)}](z) = \frac{\sin(\pi bz/2)}{\pi bz/2} B[\mathcal{D}^{(L)}](z). \tag{3.79}$$

Note that the zeros in the sine function have the effect of demoting the double poles present in $B[\mathcal{D}^{(L)}](z)$ to simple poles. As a result, the leading-b Borel transform of the $R_{e^+e^-}$ ration contains simple poles only. From these results, we can obtain resummed expressions from the observables, written in terms of (suitably regulated) exponential integral functions.

Figure 3.4 shows a diagrammatic representation of the singularities in the Borel plane, of the leading-b Borel transform of the QCD Adler D-function. There are several interesting points to note:

- The singularities can be split into two different types: those on the negative real semi-axis and those on the positive real semi-axis. The former do not pose a problem when we attempt to evaluate the Borel integral, since they do not lie along the contour of integration. By examining the path from Eqs. (3.56) and (3.57) to Eqs. (3.74) - (3.76), we can see that these singularities are generated by

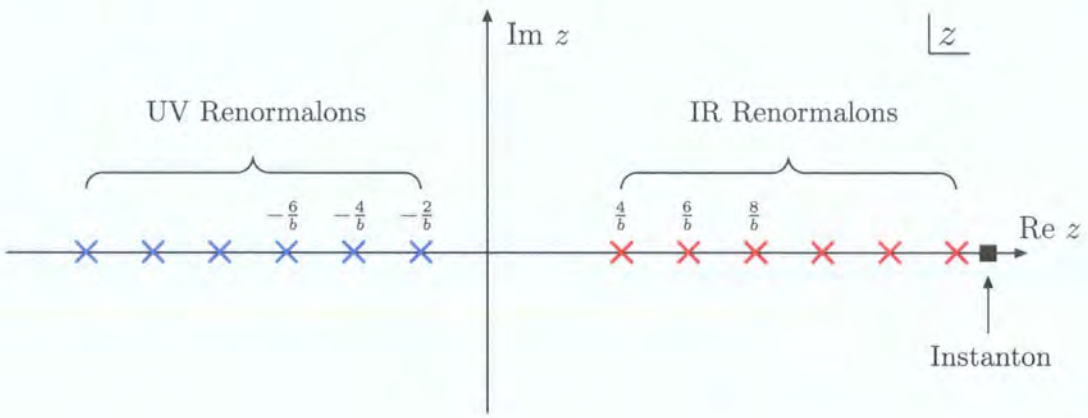


Figure 3.4: The structure of the Borel plane for the Adler D-function indicating the positions of the first six UV and IR renormalon singularities (×) and the first instanton singularity (■). The position of instanton singularities relative to the renormalon singularities will depend on the value of N_f adopted.

the UV sector of QCD, and we interpret this as signifying that perturbative QCD is well-defined in the high energy domain. Conversely, the poles on the positive real semi-axis *do* lie on the integration contour and therefore they pose a problem for Borel resummation. Such singularities mean that the Borel integral cannot be evaluated unambiguously, and this we interpret as a manifestation of the breakdown of perturbative QCD in the IR limit.

- The situation described above is reversed in the case of QED. Here, b_{QED} has the opposite sign, and so it is the UV renormalons that cause problems in QED. Correspondingly, it is the (extreme) UV region of QED which is characterized by a breakdown of perturbation theory.

- The components of Eq. (3.77) with the coefficients $A_{0,1}(n)$ are poles on the negative real semi-axis, generated by the UV behaviour of the skeleton diagram. Accordingly, we refer to the singularity at position $z = -z_n$ as the UV_n renormalon. Similarly, we refer a singularity in Eq. (3.77) at position $z = z_n$ as the IR_n renormalon.

- The absence of a renormalon singularity at position $z = z_1$ in the Borel plane is curious. We will see later, when we come to study the relationship between

renormalons and non-perturbative physics, that this corresponds directly to the absence of a dimension-2 operator in the OPE of the Adler D-function.

- In Fig. 3.4, the singularities due to instantons have also been marked, and these are at positions $z = 4n$ on the Borel plane, with $n = 1, 2, 3, \ldots$ [40, 144–146]. These singularities are caused by the existence of instanton-antiinstanton pairs. They are further away from the origin than their renormalon counterparts, and hence represent a less dominant divergent form of the perturbative coefficients. Furthermore, their positions on the Borel plane are fixed with respect to changes in N_f or b and hence they disappear in the large N_f or b limit. Curiously, for the special case of $N_f = 15$ we have $b = 1/2$ and the renormalon singularities are now positioned at $z_n = 4n$. Consequently, the positions of the renormalon singularities will now coincide with those of the instanton singularities. This occurs just before the flavour saturation of QCD ($N_f = 16$) at which point we lose asymptotic freedom.

When additional refinements such as sub-leading coefficients of the beta-function equation and scale logarithms are added, the leading-b coefficients are found to have the general form [128, 129, 145],

$$\text{IR renormalons} \quad \longleftrightarrow \quad d_n \sim C_n n! n^\gamma z_i^{-n}, \tag{3.80}$$

$$\text{UV renormalons} \quad \longleftrightarrow \quad d_n \sim C_n n! n^\gamma (-z_i)^{-n}. \tag{3.81}$$

As a result, the structure of the Borel transform is changed from that in Eq. (3.77). The single and double poles become branch points, yet the positions of these singularities remains fixed. The Borel transform have the following general form,

$$\text{IR renormalons} \quad \longleftrightarrow \quad B[\mathcal{D}^{(L)}](z) \sim \frac{\tilde{B}}{\left(1 - \frac{z}{z_i}\right)^{\gamma+1}}, \tag{3.82}$$

$$\text{UV renormalons} \quad \longleftrightarrow \quad B[\mathcal{D}^{(L)}](z) \sim \frac{\tilde{A}}{\left(1 + \frac{z}{z_i}\right)^{\gamma+1}}. \tag{3.83}$$

The leading-b approximation

It was stated previously, without justification, that the leading-b term in Eq. (3.70) is a good approximation to the exact form of the coefficients. We now give a justification of this claim.

If we adopt the values of C_A and C_F (defined in appendix A) specific to SU(N_c) QCD, then the NLO and NNLO coefficients for the Adler D-function become,

$$d_1 = -.115N_f + \left(.655N_c + \frac{.063}{N_c}\right), \tag{3.84}$$

$$d_2 = .086N_f^2 + N_f\left(-1.40N_c - \frac{.024}{N_c}\right) + \left(2.10N_c^2 - .661 - \frac{.180}{N_c^2}\right). \tag{3.85}$$

Equations (1.188) - (1.189) can also be rewritten as expansions in powers of b, using $N_f = -3b + 11C_A/2$,

$$d_1 = \left(\frac{11}{4} - 2\zeta_3\right)b + \frac{C_A}{12} - \frac{C_F}{8}, \tag{3.86}$$

$$\begin{aligned} d_2 = {} & \left(\frac{151}{18} - \frac{19}{3}\zeta_3\right)b^2 + C_A\left(\frac{31}{6} - \frac{5}{3}\zeta_3 - \frac{5}{3}\zeta_5\right)b \\ & + C_F\left(\frac{29}{32} - \frac{19}{2}\zeta_3 + 10\zeta_5\right)b + C_A^2\left(-\frac{799}{288} - \zeta_3\right) \\ & + C_AC_F\left(-\frac{827}{192} + \frac{11}{2}\zeta_3\right) + C_F^2\left(-\frac{23}{32}\right). \end{aligned} \tag{3.87}$$

If we now isolate the leading-b term in Eqs. (3.86) - (3.87) and re-expand that term in powers of N_f using Eq. (1.94), we obtain,

$$d_1^{(1)}b = .345b = -.115N_f + .634N_c, \tag{3.88}$$

$$d_2^{(2)}b^2 = .776b^2 = .086N_f^2 - .948N_fN_c + 2.61N_c^2. \tag{3.89}$$

Comparing with Eqs. (3.84) - (3.85), we see that Eqs. (3.88) - (3.89) approximate the sub-leading-N_f components of d_1 and d_2 well in both sign and magnitude. The coefficients of the $N_f^{n-r}N_c^r$ components are approximated with $\sim 20\%$ accuracy. However, it must be noted that for fixed N_f, numerical cancellations can conspire to reduce this accuracy. Hence, the success in approximating individual components of the coefficients

does not always correspond to an accurate approximation of the full coefficient. As is expected, the leading-N_f terms agree exactly.

We can draw similar conclusions about the coefficients of the GLS and unpolarized Bjorken sum rule, k_n and u_n. Rewriting Eqs. (2.80) - (2.81) and (2.85) - (2.86), for SU(N_c) QCD, we obtain,

$$k_1 = -.333N_f + \left(1.48N_c + \frac{.438}{N_c}\right), \tag{3.90}$$

$$k_2 = .177N_f^2 + N_f\left(-2.51N_c - \frac{.244}{N_c}\right) + \left(4.53N_c^2 + .686 + \frac{.008}{N_c^2}\right), \tag{3.91}$$

and

$$u_1 = -.444N_f + \left(1.84N_c + \frac{.688}{N_c}\right), \tag{3.92}$$

$$u_2 = .177N_f^2 + N_f\left(-3.111N_c - \frac{.481}{N_c}\right) + \left(5.77N_c^2 + 2.22 + \frac{.434}{N_c^2}\right). \tag{3.93}$$

In the leading-b expansion, these coefficients have the form,

$$k_1 = b + \left(\frac{C_A}{12} - \frac{7}{8}C_F\right), \tag{3.94}$$

$$\begin{aligned} k_2 = {} & b^2\left(\frac{115}{72}\right) + b\left(\frac{335}{144} + \frac{3}{2}\zeta_3 - \frac{15}{9}\zeta_5\right)C_A + b\left(-\frac{133}{288} - \frac{5}{6}\zeta_3\right)C_F \\ & + C_A^2\left(-\frac{179}{144} - \frac{11}{4}\zeta_3\right) + C_AC_F\left(-\frac{389}{192} + \frac{11}{4}\zeta_3\right) + C_F^2\left(\frac{1}{32}\right), \end{aligned} \tag{3.95}$$

and

$$u_1 = \frac{4}{3}b + \left(\frac{1}{12}C_A - \frac{11}{8}C_F\right), \tag{3.96}$$

$$\begin{aligned} u_2 = {} & b^2\left(\frac{115}{72}\right) + b\left(\frac{1705}{432} - \frac{7}{2}\zeta_3 + 5\zeta_5\right)C_A + b\left(-\frac{335}{96} + \frac{1}{2}\zeta_3\right)C_F \\ & + C_A^2\left(\frac{235}{1296} + \frac{137}{12}\zeta_3 - \frac{115}{6}\zeta_5\right) + C_AC_F\left(-\frac{2413}{192} - \frac{125}{4}\zeta_3 + \frac{95}{2}\zeta_5\right) \\ & + C_F^2\left(\frac{313}{32} + \frac{47}{2}\zeta_3 - 35\zeta_5\right). \end{aligned} \tag{3.97}$$

Expanding the leading-b term of these equations gives,

$$k_1^{(1)} b = b = -.333N_f + 1.83N_c, \tag{3.98}$$

$$k_2^{(2)} b^2 = 1.59b^2 = .177N_f^2 - 1.95N_f N + 5.37N_c^2, \tag{3.99}$$

and

$$u_1^{(1)} b = 1.33b = -.444N_f + 7.33N_c, \tag{3.100}$$

$$u_2^{(2)} b^2 = 1.59b^2 = .177N_f^2 - 1.95N_f N + 5.37N_c^2. \tag{3.101}$$

By comparing Eqs. (3.90) - (3.93) with Eqs. (3.98) - (3.101), we see that the leading-b reproduces the components of the sub-leading-b terms equally well. However, the same caveat that applied to the Adler D-function case also applies here.

In summary, as a consequence of the success of the leading-b approximation, the all-orders leading-b result of Eqs. (3.77) and the equivalent results for $B[\mathcal{K}^{(L)}](z)$ and $B[\mathcal{U}^{(L)}](z)$ can be used as effective approximations to the full perturbative expansion of $\mathcal{D}(a)$, $\mathcal{K}(a)$ and $\mathcal{U}(a)$.

3.7 Renormalons and non-perturbative corrections

The presence of IR renormalons in the Borel transforms of perturbative expansions makes their respective Borel integrals ill-defined. The best we can achieve is to evaluate these integrals using some prescription to negotiate the singularities. However, the choice of prescription is arbitrary, and hence any result obtained will have an inherent ambiguity associated with it.

We can evaluate the magnitude of this ambiguity by evaluating the difference between the Borel integrals in two 'extreme' prescriptions: that in which the contour passes just above the singularities, and that in which it passes just below it. The magnitude of this difference is then related to the residue of the singularity:

$$\Delta\mathcal{D}_{PT}(a) = \pm i\pi \,\mathrm{Res}\left(\mathrm{e}^{-z/a} \frac{1}{\left(1 - \frac{z}{z_n}\right)^{\gamma_n}}; z_n \right)$$

$$= \mp\pi i \frac{z_n^{\gamma_n}}{\Gamma(\gamma_n)} \mathrm{e}^{-z_n/a} a^{1-\gamma_n}. \tag{3.102}$$

Clearly this situation is unphysical. In order to accept perturbative QCD as a theory consistent with what we observe, the full prediction of an observable must be free from such ambiguities.

Note that this has the form of a non-perturbative contribution, reinforcing the notion that renormalons are strongly linked to non-perturbative physics. The non-perturbative nature of $\Delta\mathcal{D}_{\mathrm{PT}}$ gives us a clue as to how this problem may be resolved. Indeed, we can infer that perturbation theory can only be unambiguously defined when supplemented by non-perturbative information. Ambiguities of identical structure to $\Delta\mathcal{D}_{\mathrm{PT}}$ arise when we attempt to calculate non-perturbative corrections, again due to the necessity of having to choose a prescription. We can then assume that the ambiguities from perturbative and non-perturbative sectors cancel with each other, leaving us with an unambiguous final result. We now briefly discuss how this mechanism works.

Significant progress has been made in the relating of IR renormalons to non-perturbative corrections, see Refs. [128, 129, 145, 147, 148]. In section 1.9.1 we saw how the OPE can be used to evaluate non-perturbative corrections. The result is an expansion in powers of $1/Q^2$, with each coefficient being related to a QCD condensate. It was noted that for the Adler D-function, the first such condensate to contribute non-perturbatively is the gluon condensate, and by dimensional analysis we can determine that this is associated with a $1/Q^4$ power correction,

$$D(Q^2) = D_{\mathrm{PT}}(a) + C_{gg}(Q/M, a(M)) \frac{\langle 0|F_{\mu\nu}F^{\mu\nu}|0\rangle}{Q^4} + \mathcal{O}\left(\frac{1}{Q^6}\right), \tag{3.103}$$

where $\mathcal{O}(1/Q^6)$ represents contributions from higher dimensional condensates.

The approach to evaluating terms in Eq. (3.103) is similar to that used in section 2.4. Each condensate must be renormalized (factorized) in order to be expressed in terms of perturbative variables viz the coupling a. Hence, M in Eq. (3.103) represents the scale at which the condensates are factorized. The factorization procedure separates gluons with high momentum, which are amenable to perturbation theory, from those with low momentum, which perturbation theory cannot describe. Terms in the OPE

can then be written in terms of a scheme invariant constant containing all the non-perturbative physics, and an expression in terms of a. The scale dependence can then be determined by studying the RGE behaviour of the condensates.

A condensate of dimension $2n$ is found to have a coefficient function, $C_n(a(Q^2))$ of the form

$$C_n(a(Q^2)) \;=\; C_n\, a^{\delta_n}\,(1+\mathcal{O}(a))\,, \tag{3.104}$$

where $\delta_n = 2d/b - cz_n$, $z_n = 2n/b$, and d is the first coefficient of the gluon anomalous dimension. Non-logarithmic UV divergences [149,150] lead to an ambiguous imaginary part in the coefficient so that $C_n = C_n^{(R)} \pm iC_n^{(I)}$, mirroring the ambiguity in the perturbative part:

$$\begin{aligned}\Delta\mathcal{D}_{\rm NP} \;&=\; \pm iC_n^{(I)}\, a^{\delta_n}\left(\frac{1}{Q^2}\right)^n (1+\mathcal{O}(a)) \\ &\approx\; \pm iC_n^{(I)}\, a^{\delta_n}\, \mathrm{e}^{-z_n/a}\,(1+\mathcal{O}(a))\,.\end{aligned} \tag{3.105}$$

The ambiguities in Eqs. (3.102) and (3.105) are of identical structure. We can therefore match them and allow ambiguities to cancel between perturbative and non-perturbative components. From this we can infer a direct correspondence between IR renormalons in the Borel plane and condensates in the OPE. Specifically between the positions of Borel plane singularities and the dimension of the condensates. Therefore, each IR renormalon at position $z = z_n$ must correspond to a condensate of dimension $2n$ [128,129]. Crucially we notice that the absence of an IR renormalon at $z = z_1$ for the Adler D-function is mirrored by the absence of a dimension-2 (order $1/Q^2$) condensate contribution to the OPE.

We can now see how renormalons are interlinked with non-perturbative physics. Perturbative and non-perturbative predictions cannot exist separately in a physically consistent manner. This interdependency can be used to apply perturbative methods in order to gain knowledge of the behaviour of non-perturbative contributions.

3.8 Summary

In this chapter we presented a brief review of the large-order behaviour of perturbative QCD. We discussed Dyson's argument and described how from it we can infer a factorial form of QED perturbative coefficients. We then described two sources of factorial divergence in QCD: instantons and renormalons. We explained how these make the perturbative expansion asymptotic and how the Borel transform can be used to make sense of asymptotic series. We then described in more detail the origin of renormalon based divergence and its relation to the UV and IR sectors of QCD.

We discussed how, using exact all-orders results for the leading-N_f component of perturbative coefficients, we can obtain Borel transforms for observables in the leading-b approximation. This allows us to approximate the full form of the perturbative expansion. Finally, we discussed how the existence of renormalon singularities in the Borel plane causes the perturbative definition of an observable to become ambiguous and how only when we take into account similar ambiguities originating from the non-perturbative OPE can the perturbative representation of an observable be considered to be defined self-consistently.

Chapter 4

Infrared freezing of Euclidean QCD observables

In this chapter, we investigate the behaviour of all-orders leading-b resummations at energy scales equal to and below the Landau pole. We also study what non-perturbative information we can determine from the relationship between renormalon and operator product expansion ambiguities.

4.1 Introduction

Thanks to asymptotic freedom, fixed-order QCD perturbation theory can potentially provide accurate approximations to physical observables at suitably large energy scales, Q. However, such a perturbative description necessarily breaks down below the Landau singularity at $Q^2 = \Lambda^2$, and the infrared (IR) behaviour unavoidably involves non-perturbative effects. In fact, non-perturbative information is needed even to make sense of perturbation theory in the large Q^2 limit, since higher perturbative coefficients exhibit factorial growth, and the perturbation series is not convergent. As we saw in chapter 3, the existence of IR renormalon singularities in the Borel transform results in the perturbative predictions for observables being ambiguous. Without some mechanism to cancel these ambiguities, the self-consistency of perturbation theory would remain questionable.

Fortunately, these ambiguities are structurally the same as terms in the operator prod-

uct expansion (OPE) (in powers of Λ^2/Q^2). Hence, OPE ambiguities and Borel representation ambiguities can compensate each other, allowing the perturbative Borel and non-perturbative OPE components to be separately well-defined, once a regulation of the Borel integral (such as principal value (PV)) has been chosen [115,116].

For $Q^2 < \Lambda^2$, however, the Borel representation breaks down, mirroring the behaviour of fixed-order perturbation theory. In a recent paper, Ref. [48], which focused on the IR freezing of the Minkowskian $R_{e^+e^-}$ ratio, it was suggested that below $Q^2 = \Lambda^2$ one should use a modified Borel representation whose ambiguities come from singularities lying on the negative real semi-axis, so-called ultraviolet (UV) renormalons. We are justified in using this modified representation since, like the standard Borel representation, it recreates exactly the correct perturbative expansion when expanded in powers of a. This modified Borel representation has ambiguities which are structurally the same as a *modified expansion* in powers of Q^2/Λ^2, and once regulated both of these components can remain separately well-defined in the IR.

In this chapter we shall show that if we postulate a QCD skeleton expansion [130,151], then the leading one-chain term reproduces the standard Borel representation for $Q^2 > \Lambda^2$, and the proposed modified Borel representation for $Q^2 < \Lambda^2$. We also study the low-Q^2 behaviour of the all-orders leading-b resummations introduced in chapter 3. In particular, we are interested in their behaviour at the Landau pole ($Q^2 = \Lambda^2$) and in the 'freezing' limit ($Q^2 \to 0$). Unexpectedly, we find that the observables remain finite at the Landau pole and freeze to a well-defined limit of zero. We show that the Landau pole behaviour is due to a subtle interplay between UV and IR renormalons and that although individual renormalon contributions show Landau-type divergence, when all renormalons are summed, these divergences cancel. This behaviour is linked to the continuity and conformal symmetry of the characteristic function of the skeleton expansion.

We also investigate the compensation between ambiguities originating from renormalons and those from the OPE. Inspired by the newly discovered IR properties of the perturbative renormalon component, we are led to develop a model for non-perturbative power corrections based on demanding that ambiguities cancel each other for all Q^2. When we combine perturbative and non-perturbative components, we find that this full result shares the attractive IR properties of the original renormalon con-

tribution.

The work presented in this chapter is based on the research carried out in Ref. [152]. The material is organized as follows:

We consider the one-chain leading-b all-orders resummations of the Euclidean observables defined in sections 1.10 and 2.7.1. We introduce the skeleton expansion and discuss how this relates to the Borel representation. We show that the skeleton expansion provides a natural link between the standard Borel representation in the UV region ($Q^2 > \Lambda^2$) and a modified Borel representation in the IR region ($Q^2 < \Lambda^2$). We then investigate the IR behaviour of the observables in these representations, highlighting the attractive IR properties of their renormalon inspired (all-orders) resummed form.

Next, we consider the relation between ambiguities generated by renormalon singularities and those present in the OPE, and we determine the nature of these ambiguities in the UV and IR limits. Using the results of this, we propose a new model for power corrections which is valid for all values of Q^2. We briefly discuss the IR properties of Minkowskian observables, specifically the $R_{e^+e^-}$ ratio. We then summarize the work carried out in this chapter before finally discussing what conclusions can be drawn from it and highlighting possible future directions research on this subject may take.

4.2 QCD skeleton expansion and the Borel representation

Consider a generic Euclidean QCD observable, $\mathcal{D}(Q^2)$, having the perturbative expansion,

$$\mathcal{D}_{\rm PT}(Q^2) \;=\; a(Q^2) + \sum_{n=1}^{\infty} d_n a^{n+1}(Q^2). \tag{4.1}$$

Here $a(Q^2) \equiv \alpha_s(Q^2)/\pi$ is the renormalized coupling. Throughout this chapter we will use the one-loop approximation for the coupling,

$$a(Q^2) \;=\; \frac{2}{b \ln(Q^2/\Lambda^2)}, \tag{4.2}$$

where b is the leading beta-function coefficient in SU(3) QCD with N_f active quark flavours, and is defined in Eq. (1.94). $Q^2 \equiv -q^2 > 0$ is the single space-like energy

scale. In the $Q^2 \to \infty$ limit, asymptotic freedom ensures that $\mathcal{D}(Q^2) \to 0$. However, our interest lies in the behaviour of $\mathcal{D}(Q^2)$ at small Q^2. Specifically, is it possible that the observable remains finite at the Landau pole or freezes to a finite IR limit, $\mathcal{D}(0)$?

This is an intrinsically non-perturbative question which cannot be answered by perturbation theory alone. One has in addition, the non-perturbative contribution arising from the OPE given in Eq. (3.103). These contributions take the form of an expansion in powers of Λ^2/Q^2,

$$\mathcal{D}_{\rm NP}(Q^2) \;=\; \sum_n \mathcal{C}_n \left(\frac{\Lambda^2}{Q^2}\right)^n, \tag{4.3}$$

where $\mathcal{C}_n$ are the coefficient functions of the OPE and are presumably calculable via some non-perturbative technique. The OPE is valid for energies $Q^2 > \Lambda^2$ only. Hence, the freezing limit, if any, of $\mathcal{D}(Q^2) = \mathcal{D}_{\rm PT}(Q^2) + \mathcal{D}_{\rm NP}(Q^2)$, depends on the behaviour of *both* components as $Q^2 \to 0$.

Perturbative freezing will not arise from fixed-order perturbation theory – truncating Eq. (4.1) at finite order leads inescapably to an unphysical Landau type singularity at $Q^2 = \Lambda^2$. One needs an all-orders resummation of Eq. (4.1) in order to obtain a result that is well-behaved in the IR. Unfortunately, exact information about the coefficients is limited, at best, to calculations of d_1 and d_2 – higher-orders are unknown. All-orders information is only available in the large-N_f limit where one expands each d_n as a polynomial in N_f of order n (see Eq. (3.63)). The leading-N_f coefficient $d_n^{[n]}$ can be computed exactly to all-orders by considering the restricted set of Feynman diagrams formed by inserting the renormalon chains of Fig. 3.1 into a basic skeleton diagram [134, 153], as discussed in chapter 3. In principle one can consider more than one chain and construct a QED skeleton expansion [154].

In QCD one can replace N_f by $(33/2 - 3b)$, and obtain an expansion of d_n in powers of b,

$$d_n \;=\; d_n^{(n)} b^n + d_n^{(n-1)} b^{n-1} + \ldots + d_n^{(0)}. \tag{4.4}$$

The leading-b term $d_n^{(L)} \equiv d_n^{(n)} b^n$ can then be used to approximate d_n [143, 155, 156] as was demonstrated in section 3.6.2, and an all-orders resummation of these terms can

be performed to obtain $\mathcal{D}^{(L)}_{\rm PT}(Q^2)$,

$$\mathcal{D}^{(L)}_{\rm PT}(Q^2) \;=\; a(Q^2) + \sum_{n=1}^{\infty} d_n^{(L)} a^{n+1}(Q^2). \tag{4.5}$$

Use of the one-loop form of the coupling in Eq. (4.2) ensures that this resummed result is RS-invariant [142, 143].

This all-orders leading-b result can be written in terms of a skeleton expansion. The leading term of the skeleton expansion arises from integrating over the momentum k flowing through the renormalon chain [130, 151, 157]:

$$\mathcal{D}^{(L)}_{\rm PT}(Q^2) \;=\; \int_0^{\infty} dt\, \omega(t) a(e^C t Q^2). \tag{4.6}$$

This is equivalent to the representation of the Adler D-function given in Eq. (3.55). The kernel $\Phi(k^2/Q^2)$ has been replaced with the so-called characteristic function of the observable, $\omega(t)$, with $t \equiv k^2/Q^2$. These two functions are related by a simple transformation. The last term in this integrand comes from the summation of the chain in Fig. 3.1, over all n (see section 3.5). The identification of this with the one-chain resummation becomes apparent when we rewrite $a(tQ^2)$ as,

$$\begin{aligned} a(tQ^2) \;&=\; a(Q^2) \frac{1}{1 + a(Q^2)\frac{b \ln t}{2}} \\ &=\; a(Q^2) \sum_{n=0}^{\infty} \left(-\frac{b a(Q^2)}{2} \ln t \right)^n , \end{aligned} \tag{4.7}$$

and compare with Eq. (3.55). Here we have used the one-loop form of the coupling. The constant C depends on the subtraction procedure used to renormalize the bubble. Standard $\overline{\rm MS}$ subtraction corresponds to $C = -5/3$. From now on we shall assume $C = 0$ which corresponds to the so-called V-scheme, $\overline{\rm MS}$ subtraction with renormalization scale $\mu^2 = e^{-5/3} Q^2$. Hence, Λ in Eq. (4.2) will always refer to that in the V-scheme.

The characteristic function satisfies the normalization condition,

$$\int_0^{\infty} dt\, \omega(t) \;=\; 1, \tag{4.8}$$

which ensures the leading $a(Q^2)$ coefficient of unity in Eq. (4.1). The form of $\omega(t)$ changes at $t = 1$; it can be written as an expansion in powers of t for $t < 1$ (the IR

region) and as an expansion in inverse powers of t for $t > 0$ (the UV region), plus additional log terms. Consequently, the range of integration in Eq. (4.6) splits into an IR and a UV part,

$$\mathcal{D}^{(L)}_{\rm PT}(Q^2) \;=\; \int_0^1 dt\, \omega_{\rm IR}(t) a(tQ^2) + \int_1^\infty dt\, \omega_{\rm UV}(t) a(tQ^2). \tag{4.9}$$

The IR part corresponds to $k^2 < Q^2$, and the UV part to $k^2 > Q^2$. By making a change of variable one can transform the leading skeleton term into the familiar Borel representation of section 3.4.1. For $Q^2 > \Lambda^2$ one has the standard Borel representation (we shall explicitly write down the required changes of variable in section 4.4),

$$\mathcal{D}^{(L)}_{\rm PT}(Q^2) \;=\; \int_0^\infty dz\, e^{-z/a(Q^2)} B[\mathcal{D}^{(L)}_{\rm PT}](z). \tag{4.10}$$

Here $B[\mathcal{D}^{(L)}_{\rm PT}](z)$ is the leading-b Borel transform, defined by,

$$B[\mathcal{D}^{(L)}_{\rm PT}](z) \;=\; \sum_{n=0}^{\infty} \frac{z^n d^{(L)}_n}{n!}. \tag{4.11}$$

We saw in chapter 3 that $B[\mathcal{D}^{(L)}_{\rm PT}](z)$ contains singularities along the real z-axis. In the large-b approximation these are single and/or double poles at positions $z = z_n$ and $z = -z_n$, with $z_n \equiv 2n/b$, $n = 1, 2, 3 \ldots$, corresponding to IR_n and UV_n renormalons respectively (see section 3.6 and Fig. 3.4). The IR_n renormalons lie on the integration contour of Eq. (4.10) and therefore cause the Borel representation to be ambiguous. The difference between routeing the contour above or below the singularity yields the magnitude of the ambiguity. From Eq. (3.102) we can determine this to be of the form,

$$\Delta\mathcal{D}^{(L)}_{\rm PT} \;\sim\; \left(\frac{\Lambda^2}{Q^2}\right)^n. \tag{4.12}$$

As noted in chapter 3, this has the same form as a term in the OPE in Eq. (4.3). Thus, OPE ambiguities associated with the $(\Lambda^2/Q^2)^n$ OPE term in $\mathcal{D}_{\rm NP}(Q^2)$ can potentially cancel against the IR_n renormalon ambiguities, allowing each component separately to be well-defined [115, 116].

In practice we shall choose to take a PV definition of the integral. The IR part of the t integration in Eq. (4.9) produces the IR renormalon part of the Borel representation, and needs to be PV regulated. The second UV component produces the UV renormalons and does not require regulation. As we shall see in the next section

the standard Borel representation of Eq. (4.10) for Euclidean quantities diverges like $\ln a(Q^2)$ at $Q^2 = \Lambda^2$ for each *individual* IR_n or UV_n renormalon contribution. However, when the full set of renormalons is resummed, the $\ln a$ divergence is cancelled and a finite result is found. We shall explore this further in sections 4.3 and 4.4.

For $Q^2 < \Lambda^2$, we have $a(Q^2) < 0$, and the representation of Eq. (4.10) is invalid. The key point is that to translate from the skeleton expansion to the Borel representation, the necessary change of variables from t to z is proportional to $a(Q^2)$. Thus, if $a(Q^2)$ changes sign, then the limits of integration in z also change sign, yielding a modified Borel representation,

$$\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2) \;=\; \int_0^{-\infty} dz\, e^{-z/a(Q^2)} B[\mathcal{D}^{(L)}_{\mathrm{PT}}](z). \tag{4.13}$$

This is the modified Borel representation proposed in Ref. [48], where it was motivated as a standard Borel representation corresponding to an expansion in $|a(Q^2)| = -a(Q^2)$, since by changing variables one can write Eq. (4.13) as,

$$\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2) \;=\; -\int_0^{\infty} dz\, e^{-z/|a(Q^2)|} B[\mathcal{D}^{(L)}_{\mathrm{PT}}](-z). \tag{4.14}$$

So we see that the one-chain skeleton contribution of Eq. (4.6) is equivalent to the standard Borel representation of Eq. (4.10) for $Q^2 > \Lambda^2$, and to the modified representation of Eq. (4.13) for $Q^2 < \Lambda^2$. Note that when we substitute Eq. (4.11) into the Borel representation of Eq. (4.14) and perform successive integration by parts, it reproduces the correct form of the perturbative expansion in Eq. (4.1), for negative a.

The modified Borel representation now has a contour of integration along the *negative* real semi-axis, and so it is rendered ambiguous by the ultraviolet renormalon singularities, UV_n. Correspondingly the IR component of Eq. (4.9) is well-defined, and it is now the UV component which requires regulation. The ambiguity from routeing the contour around the singularity is now,

$$\Delta\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2) \;\sim\; \left(\frac{Q^2}{\Lambda^2}\right)^n. \tag{4.15}$$

It was suggested in Ref. [48] that the usual OPE of Eq. (4.3) breaks down for $Q^2 < \Lambda^2$, as does the associated PT Borel representation of Eq. (4.10), and should be recast and

replaced by a *modified expansion*[1] in powers of Q^2/Λ^2,

$$\mathcal{D}_{\rm NP}(Q^2) \;=\; \sum_n \tilde{C}_n \left(\frac{Q^2}{\Lambda^2}\right)^n. \tag{4.16}$$

The ambiguity associated with the $n^{\rm th}$ term in this expansion is then structurally of the same form as the ambiguity associated with the UV_n renormalon contribution in Eq. (4.13). It was further suggested in Ref. [48] that a $\tilde{C}_0$ term independent of Q^2 could arise from rearrangement of the standard OPE. This was motivated by a simple toy example. In fact, in the one-chain approximation no such term arises and both PT and NP components freeze to zero. The terms in Eq. (4.16) are then in one-to-one correspondence with the UV_n renormalon ambiguities. From its definition, the QCD skeleton expansion implies $\mathcal{D}_{\rm PT}(0) = 0$ in the $Q^2 \to 0$ limit. For the one-chain term in Eq. (4.6) this follows simply because as $Q^2 \to 0$ the integrand vanishes everywhere in the range of integration, since $a(tQ^2) \to 0$ for any given t. Higher multiple chain terms will contain products of the form $a(t_1Q^2)a(t_2Q^2)\ldots$ in the integrand and will similarly vanish. Hence, this implies that in the IR limit, $\mathcal{D}_{\rm NP}(Q^2)$ behaves as,

$$\mathcal{D}_{\rm NP}(Q^2) \;\approx\; k\left(\frac{Q^2}{\Lambda^2}\right)^{n_0}, \tag{4.17}$$

where UV_{n_0} is the UV renormalon singularity nearest to the origin in the Borel plane.

We should note that the modified Borel representation, its IR behaviour and its connection with UV renormalons, has also been discussed in Ref. [131]. In Appendix B of that paper the IR freezing behaviour of the Adler D-function, $D(Q^2)$, was discussed. From general arguments of non-perturbative spontaneous chiral symmetry breaking in the limit of a large number of colours, N_c, it was concluded that in the $Q^2 \to 0$ limit, $D(Q^2)$ goes to zero like,

$$D(Q^2) \;\sim\; \frac{Q^2}{M^2}. \tag{4.18}$$

Here, M is the mass of a one-meson state, with these states remaining massive in the chiral limit. A similar result is also obtained in Ref. [158]. Since UV_1 is the singularity nearest the origin for the Adler D-function, then $n_0 = 1$, and therefore the freezing expectation is indeed consistent with Eq. (4.17). Note that strictly, the leading

[1] This could be considered to be the IR equivalent of the OPE – applicable to $Q^2 < \Lambda^2$. However, the OPE can only be defined self-consistently for $Q^2 > \Lambda^2$ and therefore any terminology associated with the OPE is inappropriate. We therefore refer to this as the 'modified expansion'.

behaviour as $Q^2 \to 0$ is the logarithmic freezing to zero of $a(Q^2)$ contributed by the PT component. It is the non-perturbative effects which reflect the UV renormalon structure.

4.3 The Q^2-dependence of Euclidean observables

The four Euclidean observables we shall consider are defined in sections 1.10 and 2.7.1. Here, we briefly summarize their origins and definitions.

The QCD vacuum polarization function, $\Pi(Q^2)$, is the correlator of two vector currents in the Euclidean region,

$$(q_\mu q_\nu - g_{\mu\nu}q^2)\Pi(Q^2) \;=\; 16\pi^2 i \int d^4x\, e^{iq.x}\langle 0|T[j_\mu(x)j_\nu(0)]|0\rangle, \qquad (4.19)$$

and the leading-N_f component of $\Pi(Q^2)$ can be calculated from the diagrams in Fig. 3.3. The Adler D-function, $D(Q^2)$, is then defined via the logarithmic derivative of $\Pi(Q^2)$,

$$D(Q^2) \;=\; -\frac{3}{4}Q^2\frac{d}{dQ^2}\Pi(Q^2). \qquad (4.20)$$

This can be split into the parton model result and QCD corrections, $\mathcal{D}(Q^2)$,

$$D(Q^2) \;=\; N_c\sum_f Q_f^2\left(1+\frac{3}{4}C_F\mathcal{D}(Q^2)\right), \qquad (4.21)$$

where N_c is the number of colours, Q_f is the charge of quark flavour f, and C_F is defined in appendix A. Here, $\mathcal{D}(Q^2) = \mathcal{D}_{\rm PT}(Q^2) + \mathcal{D}_{\rm NP}(Q^2)$, with the two components defined as in Eqs. (4.1) and (4.3). The polarized Bjorken (pBj) [89,90] and GLS [81] sum rules were defined in section 2.7.1. They too can be separated into a parton model result and QCD corrections:

$$K_{pBj}(Q^2) \;=\; \frac{1}{3}\left|\frac{g_A}{g_V}\right|\left(1-\frac{3}{4}C_F\mathcal{K}(Q^2)\right), \qquad (4.22)$$

$$K_{GLS}(Q^2) \;=\; 3\left(1-\frac{3}{4}C_F\mathcal{K}(Q^2)\right). \qquad (4.23)$$

Here, we have neglected contributions due to "light-by-light" diagrams – which when

omitted render the perturbative corrections to $K_{GLS}(Q^2)$ and $K_{pBj}(Q^2)$ identical. Finally, the unpolarized Bjorken sum rule (uBj) [85] (also defined in section 2.7.1) can be written,

$$U_{uBj} = \left(1 - \frac{1}{2}C_F\mathcal{U}(Q^2)\right). \tag{4.24}$$

In these expressions, the QCD corrections $\mathcal{K}(Q^2)$ and $\mathcal{U}(Q^2)$ are split into PT and NP components, as for $\mathcal{D}(Q^2)$. The leading-N_f contributions to these three sum rules can be calculated from the diagrams in Fig. 4.1. These large-N_f results can then be used to compute leading-b all-orders resummations for these observables, $\mathcal{D}^{(L)}_{\rm PT}(Q^2)$, $\mathcal{K}^{(L)}_{\rm PT}(Q^2)$ and $\mathcal{U}^{(L)}_{\rm PT}(Q^2)$, as was described in Section 3.6.

The leading-b Borel transforms of the above observables are well known. For $\mathcal{D}^{(L)}_{\rm PT}(Q^2)$, the Borel transform can be found in Ref. [143], and has the form,[2]

$$\begin{aligned} B[\mathcal{D}^{(L)}_{\rm PT}](z) &= \sum_{n=1}^{\infty} \frac{A_0(n) - A_1(n)z_n}{\left(1+\frac{z}{z_n}\right)^2} + \frac{A_1(n)z_n}{\left(1+\frac{z}{z_n}\right)} \\ &+ \sum_{n=1}^{\infty} \frac{B_0(n) + B_1(n)z_n}{\left(1-\frac{z}{z_n}\right)^2} - \frac{B_1(n)z_n}{\left(1-\frac{z}{z_n}\right)}. \end{aligned} \tag{4.25}$$

Here,

$$\begin{aligned} A_0(n) &= \frac{8}{3}\frac{(-1)^{n+1}(3n^2+6n+2)}{n^2(n+1)^2(n+2)^2}, & A_1(n) &= \frac{8}{3}\frac{b(-1)^{n+1}(n+\frac{3}{2})}{n^2(n+1)^2(n+2)^2}, \\ B_0(1) &= 0, & B_1(1) &= 0, \\ B_0(2) &= 1, & B_1(2) &= -\frac{b}{4}, \\ B_0(n) &= -A_0(-n), & B_1(n) &= -A_1(-n), \qquad \text{for } n \geq 3. \end{aligned} \tag{4.26}$$

The Borel transforms of $\mathcal{K}^{(L)}_{\rm PT}(Q^2)$ and $\mathcal{U}^{(L)}_{\rm PT}(Q^2)$ can be found in Refs. [135,159] respectively. They have a much simpler structure than that of the Adler D-function since they arise from the insertion of a renormalon chain into a tree-level diagram, rather than into a quark loop, as can be seen by comparing Fig. 4.1 with Fig. 3.3. Consequently,

[2] Although we have already given the Borel transform of the Adler D-function in chapter 3, we restate it here for ease of reference.

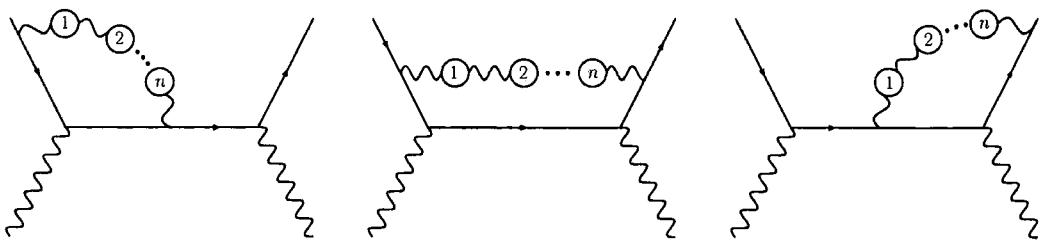


Figure 4.1: Leading leading-N_f contributions to the DIS sum rules of Eqs. (4.22) - (4.24) at n^{th} order in perturbation theory.

there are only a finite number of single poles and no double poles, and we can write out their Borel transforms explicitly:

$$B[\mathcal{K}^{(L)}_{\text{PT}}](z) = \frac{4/9}{\left(1+\frac{z}{z_1}\right)} - \frac{1/18}{\left(1+\frac{z}{z_2}\right)} + \frac{8/9}{\left(1-\frac{z}{z_1}\right)} - \frac{5/18}{\left(1-\frac{z}{z_2}\right)}, \tag{4.27}$$

and

$$B[\mathcal{U}^{(L)}_{\text{PT}}](z) = \frac{1/6}{\left(1+\frac{z}{z_2}\right)} + \frac{4/3}{\left(1-\frac{z}{z_1}\right)} - \frac{1/2}{\left(1-\frac{z}{z_2}\right)}. \tag{4.28}$$

As noted in section 3.6.2 the leading-b approximations for the NLO and NNLO coefficients for these observables are in reasonable agreement with the known exact coefficients [143, 159].

We can now evaluate the Borel integral of Eq. (4.10) to obtain $\mathcal{D}^{(L)}_{\text{PT}}(Q^2)$, $\mathcal{K}^{(L)}_{\text{PT}}(Q^2)$ and $\mathcal{U}^{(L)}_{\text{PT}}(Q^2)$. Using the integrals,

$$\int_0^\infty dz \frac{e^{-z/a}}{(1+z/z_n)} = -z_n e^{z_n/a}\text{Ei}(-z_n/a), \tag{4.29}$$

$$\int_0^\infty dz \frac{e^{-z/a}}{(1+z/z_n)^2} = z_n\left[1+\frac{z_n}{a}e^{z_n/a}\text{Ei}(-z_n/a)\right], \tag{4.30}$$

the following resummed expressions are obtained,

$$\mathcal{D}^{(L)}_{\rm PT}(Q^2) = \sum_{n=1}^{\infty} z_n \left\{ e^{z_n/a(Q^2)} {\rm Ei}\left(-\frac{z_n}{a(Q^2)}\right) \left[\frac{z_n}{a(Q^2)}\left(A_0(n) - z_n A_1(n)\right) - z_n A_1(n)\right] + \left(A_0(n) - z_n A_1(n)\right) \right\}$$
$$+ \sum_{n=1}^{\infty} z_n \left\{ e^{-z_n/a(Q^2)} {\rm Ei}\left(\frac{z_n}{a(Q^2)}\right) \left[\frac{z_n}{a(Q^2)}\left(B_0(n) + z_n B_1(n)\right) - z_n B_1(n)\right] - \left(B_0(n) + z_n B_1(n)\right) \right\}, \quad (4.31)$$

$$\mathcal{K}^{(L)}_{\rm PT}(Q^2) = \frac{1}{9b}\left[-8e^{z_1/a(Q^2)} {\rm Ei}\left(-\frac{z_1}{a(Q^2)}\right) + 2e^{z_2/a(Q^2)} {\rm Ei}\left(-\frac{z_2}{a(Q^2)}\right) + 16 e^{-z_1/a(Q^2)} {\rm Ei}\left(\frac{z_1}{a(Q^2)}\right) - 10 e^{-z_2/a(Q^2)} {\rm Ei}\left(\frac{z_2}{a(Q^2)}\right)\right], \quad (4.32)$$

$$\mathcal{U}^{(L)}_{\rm PT}(Q^2) = \frac{1}{3b}\left[8e^{-z_1/a(Q^2)} {\rm Ei}\left(\frac{z_1}{a(Q^2)}\right) - 6 e^{-z_2/a(Q^2)} {\rm Ei}\left(\frac{z_2}{a(Q^2)}\right) - 2e^{z_2/a(Q^2)} {\rm Ei}\left(-\frac{z_2}{a(Q^2)}\right)\right], \quad (4.33)$$

where ${\rm Ei}(x)$ is the exponential integral function (defined in section B.3) and for $x > 0$ we must take the PV of the integral in its definition. ${\rm Ei}(x)$ has the expansion,

$${\rm Ei}(x) = \ln|x| + \gamma_{\rm E} + \mathcal{O}(x), \quad (4.34)$$

for small x, where $\gamma_{\rm E} = 0.57721\ldots$ is the Euler constant.

A crucial point is that the above expressions for the Q^2-dependence apply at *all* values of Q^2. For $Q^2 < \Lambda^2$ the modified Borel representation (written as an ordinary Borel representation for an expansion in powers of $|a|$, as in Eq. (4.14)) corresponds to changing $a(Q^2) \to -a(Q^2)$, $z_n \to -z_n$, and adding an overall minus sign in Eqs. (4.31) - (4.33). One can easily see that these equations are invariant under these changes. Note

that in Eq. (4.31) one needs to change $A_1 \to -A_1$ and $B_1 \to -B_1$, since they contain a hidden z_n factor in their definitions. Also, in Eqs. (4.32) and (4.33) the prefactor proportional to $1/b$ also needs to change sign since it has been factorized from z_1, z_2.

The expressions in Eqs. (4.31) - (4.33) are the leading-b components (i.e. $d_n^{(n)} b^n a^{n+1}$) of the perturbative expansions, resummed to all-orders. Although, one might argue that these expressions omit an infinite set of terms in the full perturbative expansion, namely those sub-leading in b, we must remember so too do fixed-order predictions. In fixed-order perturbation theory, we omit an infinite subset of terms which are suppressed by powers of the coupling but which we know to be factorially divergent at higher orders. In all-orders perturbation theory, we include the dominant divergent components of the coefficients (the leading-b term) and omit an infinite subset of terms which, although they do not fully include the known leading-α_s terms, we know to be less divergent, and furthermore, appear to be naturally suppressed relative to the leading-b component (see section 3.6.2).

4.3.1 Infrared behaviour

We now turn our attention to the behaviour of the above resummed expressions at the Landau pole. The $\mathrm{Ei}(z_n/a(Q^2))$ functions exhibit a logarithmic divergence as their argument goes to zero, and so it would appear that one does not obtain a finite result at $Q^2 = \Lambda^2$. Using Eq. (4.34), we have,

$$\begin{aligned}
\mathrm{Ei}\left[\frac{2n}{ba(Q^2)}\right] &= \mathrm{Ei}\left[n\log(Q^2/\Lambda^2)\right] \\
&= \mathrm{Ei}\left[n\left(\frac{Q^2}{\Lambda^2}-1\right)+\ldots\right] \\
&= \gamma_{\mathrm{E}} + \ln\left[n\left(\frac{Q^2}{\Lambda^2}-1\right)\right] + \mathcal{O}\left(\frac{Q^2}{\Lambda^2}-1\right), \qquad (4.35)
\end{aligned}$$

for $\Lambda^2 \approx Q^2$. Note that the only terms in Eqs. (4.31) - (4.33) which could possibly contribute to any divergence are the $e^{\pm z_n/a}\mathrm{Ei}(\mp z_n/a)$ terms, and these are present in both the UV and IR components. All other terms vanish or are finite. As can be seen from Eqs. (4.29) and (4.30), these terms are generated exclusively by the single pole terms in the Borel transform. The double pole terms only generate finite contributions

at $Q^2 = \Lambda^2$.

Using Eq. (4.35) we obtain the $Q^2 \to \Lambda^2$ limit of $\mathcal{D}_{\rm PT}^{(L)}(Q^2)$,

$$\begin{aligned}\mathcal{D}_{\rm PT}^{(L)}(Q^2) &= -\sum_{n=1}^{\infty} z_n^2[A_1(n)+B_1(n)]\ln\left[n\left(\frac{Q^2}{\Lambda^2}-1\right)\right] \\ &+ \sum_{n=1}^{\infty}[z_n^2(1+\gamma_{\rm E})(-A_1(n)-B_1(n))+z_n(A_0(n)-B_0(n))] \\ &+ \mathcal{O}\left(\frac{Q^2}{\Lambda^2}-1\right). \end{aligned} \tag{4.36}$$

So the coefficient of the divergent log term in $\mathcal{D}_{\rm PT}^{(L)}(Q^2)$ is,

$$-\sum_{n=1}^{\infty} z_n^2[A_1(n)+B_1(n)], \tag{4.37}$$

and for $\mathcal{K}_{\rm PT}^{(L)}(Q^2)$ and $\mathcal{U}_{\rm PT}^{(L)}(Q^2)$ the equivalent coefficients are $(-8+2+16-10=0)$ and $(8-6-2=0)$, respectively. Cancellation clearly occurs in the cases of $\mathcal{K}_{\rm PT}^{(L)}(Q^2)$ and $\mathcal{U}_{\rm PT}^{(L)}(Q^2)$ and in the case of $\mathcal{D}_{\rm PT}^{(L)}(Q^2)$ the previously unnoticed relation

$$z_{n+3}^2 B_1(n+3) = -z_n^2 A_1(n), \tag{4.38}$$

ensures that the potential logarithmic divergence in Eq. (4.31) cancels, and that $\mathcal{D}_{\rm PT}^{(L)}(\Lambda^2)$ is finite,

$$\sum_{n=1}^{\infty} z_n^2[A_1(n)+B_1(n)] = 0. \tag{4.39}$$

A similar relation,

$$A_0(n) = -B_0(n+2), \tag{4.40}$$

was noted in Ref. [143]. We shall show in the next section that the relations of Eqs. (4.38) and (4.40) are underwritten by the continuity of the skeleton expansion characteristic function $\omega_\Pi(t)$ and its first derivative at $t=1$. The form of the perturbative corrections, $\mathcal{D}_{\rm PT}^{(L)}(Q^2)$, $\mathcal{K}_{\rm PT}^{(L)}(Q^2)$ and $\mathcal{U}_{\rm PT}^{(L)}(Q^2)$, are shown in Fig. 4.2.

Although we have shown that when summed to infinity, Eq. (4.31) is finite at $Q^2 = \Lambda^2$,

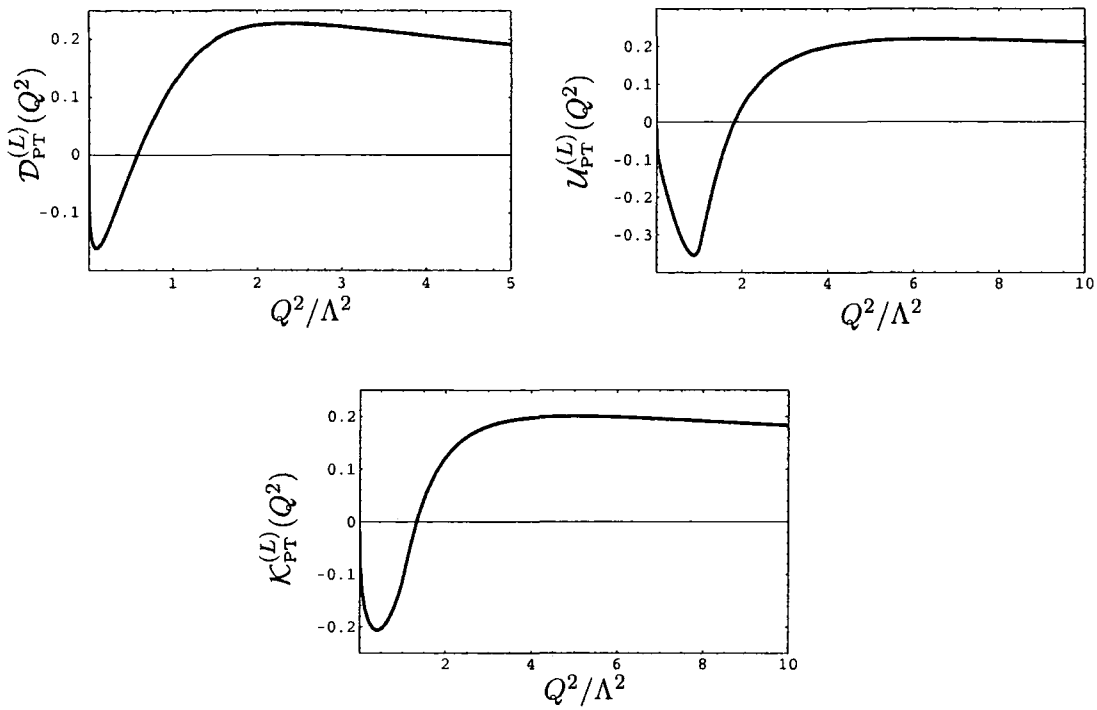


Figure 4.2: Q^2-dependence of the leading-b perturbative corrections to the observables in Eqs. (4.21) - (4.24), resummed to all-orders in the leading-b approximation.

obviously we can only plot the expression including a finite number of terms in the n sum. The expression can remain finite, however, if we sum the UV renormalons to finite $n = N$, and the IR renormalons to $n = N + 3$. In this case the relation of Eq. (4.38) will ensure that the divergent terms cancel. In Fig. 4.2 we have taken $N = 50$ and assumed $N_f = 0$ quark flavours, thus avoiding the need to match at quark flavour thresholds (since we are only interested here in the form of the freezing behaviour, not in a phenomenological analysis).

The plots in Fig. 4.2 demonstrate two important points about the Euclidean quantities we are considering. Firstly, they are finite at $Q^2 = \Lambda^2$, and secondly they change sign just above or below this point, before freezing to zero at $Q^2 = 0$ (as noted in section 4.2). Note that although these perturbative corrections become negative, crucially the full observable (i.e. with the parton model result included) remains positive at all values of Q^2.

The relation of Eq. (4.38) simplifies the expression for the finite part of Eq. (4.36),

hence it becomes,

$$\begin{aligned}\mathcal{D}^{(L)}_{\rm PT}(Q^2=\Lambda^2) &= \sum_{n=1}^{\infty} z_n[A_0(n)-B_0(n)] - \sum_{n=1}^{\infty} z_n^2[A_1(n)+B_1(n)]\ln n \\ &= \frac{0.679938\ldots}{b}. \end{aligned} \tag{4.41}$$

The values $\mathcal{K}^{(L)}_{\rm PT}(Q^2=\Lambda^2)$ and $\mathcal{U}^{(L)}_{\rm PT}(Q^2=\Lambda^2)$ are given by a formula identical to Eq. (4.41), but using values of $A_{0,1}(n)$ and $B_{0,1}(n)$ appropriate to $\mathcal{K}(Q^2)$ and $\mathcal{U}(Q^2)$. Although we have not given these values explicitly, they are of a much simpler form than in the case of $\mathcal{D}(Q^2)$, and they can easily be deduced by comparing Eqs. (4.27) and (4.28) with Eq. (4.25). From this we obtain,

$$\begin{aligned}\mathcal{K}^{(L)}_{\rm PT}(Q^2=\Lambda^2) &= -\frac{8}{9b}\ln 2, \\ \mathcal{U}^{(L)}_{\rm PT}(Q^2=\Lambda^2) &= -\frac{8}{3b}\ln 2. \end{aligned} \tag{4.42}$$

The fact that these resummed expressions are finite at the Landau pole is quite remarkable, and is in stark contrast to the disastrous behaviour of fixed-order expressions at this energy scale. Summing the leading-b component of the perturbative expansion to all orders would appear to cancel the logarithmic pole in the one-loop coupling,

$$\begin{aligned}\mathcal{D}^{(L)}_{\rm PT}(Q^2=\Lambda^2) &= \lim_{Q^2\to\Lambda^2}\sum_{n=0}^{\infty} d_n^{(n)} b^n \left(\frac{2}{b\ln(Q^2/\Lambda^2)}\right)^{n+1} \\ &= \frac{0.679938\ldots}{b}. \end{aligned} \tag{4.43}$$

Furthermore, from the results in chapter 3 we know that the above result is asymptotic to the exact form of the observable, in the large-b limit. It is also interesting that this cancellation occurs because of the conformal symmetry and other relations between the IR and UV sectors of QCD, expressed in terms of the characteristic function.

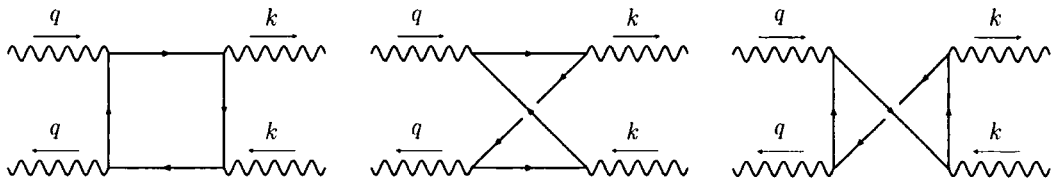


Figure 4.3: Light-by-light scattering diagrams, used to calculate $\omega_\Pi(t)$.

4.4 The skeleton expansion and Borel representations for the Adler D-function

In this section we show how the skeleton expansion relates to the Borel representation and how from this relationship we can uncover the origin of the Landau pole behaviour of Euclidean observables. We begin with the one-chain skeleton expansion result for the vacuum polarization function $\Pi(Q^2)$ defined in Eqs. (1.183) and (4.19),

$$\Pi(Q^2) \;=\; \int_0^\infty dt\, \omega_\Pi(t) a(tQ^2), \tag{4.44}$$

where in this case the characteristic function $\omega_\Pi(t)$ is given by,

$$\omega_\Pi(t) \;=\; \frac{4}{3}\begin{cases} t\,\Xi(t) & t \le 1 \;\leftrightarrow\; \text{IR} \\ \frac{1}{t}\,\Xi\left(\frac{1}{t}\right) & t \ge 1 \;\leftrightarrow\; \text{UV} \end{cases} \tag{4.45}$$

and can be obtained from the classic QED work of Ref. [132] by simply including appropriate colour factors. In this language it is related to the Bethe-Salpeter kernel for the scattering of light-by-light, and is the first term in a well-defined QED skeleton expansion [154]. In contrast, the existence of the QCD skeleton expansion beyond leading-b is more problematic [157].

The diagrams relevant to this kernel are shown in Fig. 4.3. It is easy to see how, by connecting the ends of the renormalon chain in Fig. 3.1 to the ends of the momentum-k external propagators in Fig. 4.3, one can reproduce the topology of the diagrams in Fig. 3.3.

The exact form of $\Xi(t)$ can be found in Ref. [132],

$$\Xi(t) \equiv \frac{4}{3t}\left\{1 - \ln t + \left(\frac{5}{2} - \frac{3}{2}\ln t\right)t + \frac{(1+t)^2}{t}\left[\mathrm{Li}_2(-t) + \ln t \ln(1+t)\right]\right\}, \tag{4.46}$$

where $\mathrm{Li}_2(x)$ is the dilogarithm function, defined in Eq. (B.13). Although we define $\omega_\Pi(t)$ separately in the IR and UV domains, the two regions are related by the conformal symmetry $t \leftrightarrow \frac{1}{t}$.

The Adler D-function (related to $\Pi(Q^2)$ through Eq. (4.20)) will have the following one-chain skeleton expansion term with characteristic function $\omega_\mathcal{D}(t)$,

$$\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2) = \int_0^\infty dt\, \omega_\mathcal{D}(t) a(tQ^2). \tag{4.47}$$

$\omega_\mathcal{D}(t)$ can be obtained from $\omega_\Pi(t)$ by performing the differentiation of Eq. (4.20) on Eq. (4.44) and then performing integration by parts on the resulting expression:

$$\begin{aligned}
\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2) &= -\frac{3}{4}Q^2\frac{d}{dQ^2}\int_0^\infty dt\, \omega_\Pi(t) t\left(\frac{a(tQ^2)}{t}\right) \\
&= +\frac{3}{2b}Q^2\frac{d}{dQ^2}\int_0^\infty dt \frac{d}{dt}\Big[\omega_\Pi(t)t\Big]\ln[a(tQ^2)] \\
&= \frac{3}{4}\int_0^\infty dt\Big[\omega_\Pi(t) + t\frac{d}{dt}\omega_\Pi(t)\Big]a(tQ^2).
\end{aligned} \tag{4.48}$$

The transformation from $\Pi(Q^2)$ to $\mathcal{D}(Q^2)$ therefore induces a transformation in $\omega_\Pi(t)$ of,

$$\begin{aligned}
\Pi(Q^2) &\to Q^2\frac{d}{dQ^2}\Pi(Q^2) = -\frac{4}{3}\mathcal{D}(Q^2), \\
\Rightarrow \quad \omega_\Pi(t) &\to \omega_\Pi(t) + t\frac{d}{dt}\omega_\Pi(t) = \frac{4}{3}\omega_\mathcal{D}(t).
\end{aligned} \tag{4.49}$$

This transformation spoils the conformal symmetry present in $\omega_\Pi(t)$. Indeed, the expressions for $\omega_\mathcal{D}(t)$ in the UV and IR regions are slightly more complicated,

$$\omega^{\mathrm{IR}}_\mathcal{D}(t) = \frac{8}{3}\left\{\left(\frac{7}{4} - \ln t\right)t + (1+t)\Big[\mathrm{Li}_2(-t) + \ln t \ln(1+t)\Big]\right\}, \tag{4.50}$$

$$\omega_{\mathcal{D}}^{\rm UV}(t) = \frac{8}{3}\Bigg\{1+\ln t+\left(\frac{3}{4}+\frac{1}{2}\ln t\right)\frac{1}{t} + (1+t)\Big[{\rm Li}_2(-t^{-1})-\ln t\ln(1+t^{-1})\Big]\Bigg\}. \quad (4.51)$$

However, a partial symmetry remains in $\omega_{\mathcal{D}}(t)$ and this will be elucidated upon in the following discussion.

We shall now convert the skeleton expansion form into the Borel representations of Eqs. (4.10) and (4.13) by making a change of variables. To achieve this, it is necessary to write $\omega_{\Pi}(t)$ as an expansion in powers of t. This yields expressions in both the IR and UV regions comprising of an expansion plus an expansion times a logarithm,

$$\omega_{\Pi}^{\rm IR}(t) = \frac{4}{3}\left(\sum_{n=1}^{\infty}\xi_n t^n+\ln t\sum_{n=2}^{\infty}\hat{\xi}_n t^n\right). \quad (4.52)$$

The conformal symmetry expressed in Eq. (4.45) means that the UV part can also be written in terms of the same coefficients, ξ_n and $\hat{\xi}_n$,

$$\omega_{\Pi}^{\rm UV}(t) = \frac{4}{3}\left(\sum_{n=1}^{\infty}\xi_n t^{-n}-\ln t\sum_{n=2}^{\infty}\hat{\xi}_n t^{-n}\right). \quad (4.53)$$

From Eq. (4.46), ξ_n and $\hat{\xi}_n$ are found to be,

$$\xi_{n>1} = \frac{4}{3}\frac{(2-6n^2)(-1)^n}{(n-1)^2n^2(n+1)^2}, \qquad \hat{\xi}_{n>1} = \frac{4}{3}\frac{2(-1)^n}{(n-1)n(n+1)},$$
$$\xi_1 = 1, \qquad \hat{\xi}_1 = 0. \quad (4.54)$$

Performing the transformation in Eq. (4.49) allows us to write $\omega_{\mathcal{D}}(t)$ as a similar expansion,

$$\omega_{\mathcal{D}}^{\rm IR}(t) = \sum_{n=1}^{\infty}[\xi_n(1+n)+\hat{\xi}_n]t^n+\ln t\sum_{n=2}^{\infty}\hat{\xi}_n(n+1)t^n, \quad (4.55)$$
$$\omega_{\mathcal{D}}^{\rm UV}(t) = \sum_{n=1}^{\infty}[\xi_n(1-n)-\hat{\xi}_n]t^{-n}+\ln t\sum_{n=2}^{\infty}\hat{\xi}_n(n-1)t^{-n}. \quad (4.56)$$

Using the expansions of Eqs. (4.55) and (4.56) we can now represent $\mathcal{D}_{\rm PT}^{(L)}(Q^2)$ in terms of a Borel integral. We take $\mathcal{D}_{\rm PT}^{(L)}(Q^2)$ expressed in terms of $\omega_{\mathcal{D}}(t)$ and then split the

integral into IR and UV regions,

$$\begin{aligned}
\mathcal{D}_{\rm PT}^{(L)}(Q^2) &= \int_0^\infty dt\,\omega_{\mathcal{D}}(t)a(tQ^2) \\
&= \sum_{k=0}^{\infty} a(Q^2)\int_0^1 dt\,\omega_{\mathcal{D}}^{\rm IR}(t)\left(-\frac{ba(Q^2)}{2}\ln t\right)^k \\
&+ \sum_{k=0}^{\infty} a(Q^2)\int_1^\infty dt\,\omega_{\mathcal{D}}^{\rm UV}(t)\left(-\frac{ba(Q^2)}{2}\ln t\right)^k \\
&= a(Q^2)\sum_{k=0}^{\infty}\left(-\frac{ba(Q^2)}{2}\right)^k \\
&\left[\int_0^1 dt\left(\sum_{n=1}^{\infty}[\xi_n(1+n)+\hat{\xi}_n](t)^n + \ln t\sum_{n=2}^{\infty}\hat{\xi}_n(n+1)(t)^n\right)(\ln t)^k\right. \\
&+ \left.\int_1^\infty dt\left(\sum_{n=1}^{\infty}[\xi_n(1-n)-\hat{\xi}_n](t)^{-n} + \ln t\sum_{n=2}^{\infty}\hat{\xi}_n(n-1)(t)^{-n}\right)(\ln t)^k\right].
\end{aligned} \tag{4.57}$$

Here we have used,

$$a(xy) = a(y)\sum_{k=0}^{\infty}\left(-\frac{ba(y)}{2}\ln x\right)^k. \tag{4.58}$$

We note that $[\xi_n(1-n)-\hat{\xi}_n]=0$ for $n=1$, which allows us to omit this term from Eq. (4.57). This expression may be transformed into a Borel integral of the form of Eq. (4.10) by changes of variables and integration by parts. We use the change of variables $z=-a(Q^2)(n+1)\ln t$ and $z=a(Q^2)(n-1)\ln t$ for IR and UV parts of Eq. (4.57), respectively. Integration by parts is necessary for the integrals with an extra $\ln t$ term. For $Q^2>\Lambda^2$, we have $a(Q^2)>0$, and we then obtain the standard Borel representation, of Eq. (4.10),

$$\mathcal{D}_{\rm PT}^{(L)}(Q^2) = \int_0^\infty dz\, \mathrm{e}^{\frac{-z}{a(Q^2)}} \left[\sum_{n=1}^{\infty} \frac{[\xi_n(1+n)+\hat{\xi}_n]}{n+1} \frac{1}{\left(1-\frac{bz}{2(n+1)}\right)} - \sum_{n=2}^{\infty} \frac{\hat{\xi}_n(n+1)}{(n+1)^2} \frac{1}{\left(1-\frac{bz}{2(n+1)}\right)^2} \right]$$
$$+ \int_0^\infty dz\, \mathrm{e}^{\frac{-z}{a(Q^2)}} \left[\sum_{n=2}^{\infty} \frac{[\xi_n(1-n)-\hat{\xi}_n]}{n-1} \frac{1}{\left(1+\frac{bz}{2(n-1)}\right)} + \sum_{n=2}^{\infty} \frac{\hat{\xi}_n(n-1)}{(n-1)^2} \frac{1}{\left(1+\frac{bz}{2(n-1)}\right)^2} \right]. \quad (4.59)$$

For $Q^2 < \Lambda^2$, $a(Q^2) < 0$, we obtain the modified Borel representation of Eq. (4.13), in which the upper limit in z is $-\infty$.

Having obtained the Borel transform we can now make contact with Eq. (4.25), and this allows us to make the following identifications:

$$\frac{\xi_n(1+n)+\hat{\xi}_n}{n+1} = -B_1(n+1)z_{n+1}, \qquad n \geq 1 \quad (4.60)$$

$$\frac{\xi_n(1-n)-\hat{\xi}_n}{n-1} = A_1(n-1)z_{n-1}, \qquad n \geq 2 \quad (4.61)$$

for the single pole residues and,

$$-\frac{\hat{\xi}_n(n+1)}{(n+1)^2} = B_0(n+1) + B_1(n+1)z_{n+1}, \qquad n \geq 2 \quad (4.62)$$

$$\frac{\hat{\xi}_n(n-1)}{(n-1)^2} = A_0(n-1) - A_1(n-1)z_{n-1}, \qquad n \geq 2 \quad (4.63)$$

for the double pole residues. Substituting the form of ξ_n and $\hat{\xi}_n$ given by Eq. (4.54), and comparing with Eq. (4.26), verifies the above equations.

Equations (4.60) - (4.63) can be used to rewrite the $\omega_{\mathcal{D}}^{\rm IR}(t)$ and $\omega_{\mathcal{D}}^{\rm UV}(t)$ expansions of

Eqs. (4.55) and (4.56) in terms of the $A_0(n)$, $A_1(n)$, and $B_0(n)$, $B_1(n)$ renormalon residues. One finds,

$$\omega_{\mathcal{D}}^{\rm IR}(t) = \frac{b}{2}\sum_{n=1}^{\infty} -z_{n+1}^2 B_1(n+1)t^n - \ln t \sum_{n=2}^{\infty}(n+1)^2[B_0(n+1)+z_{n+1}B_1(n+1)]t^n, \quad (4.64)$$

$$\omega_{\mathcal{D}}^{\rm UV}(t) = \frac{b}{2}\sum_{n=1}^{\infty} z_{n-1}^2 A_1(n-1)t^{-n} + \ln t \sum_{n=2}^{\infty}(n-1)^2[A_0(n-1)-z_{n-1}A_1(n-1)]t^{-n}. \quad (4.65)$$

The discontinuity at $t=1$ is then found to be,

$$\omega_{\mathcal{D}}^{\rm UV}(1)-\omega_{\mathcal{D}}^{\rm IR}(1) = \frac{b}{2}\sum_{n=1}^{\infty} z_n^2[A_1(n)+B_1(n)], \quad (4.66)$$

which vanishes using Eq. (4.39). In the language of ξ_n and $\hat{\xi}_n$ coefficients, Eq. (4.66) is equivalent to,

$$-2\sum_{n=1}^{\infty}\left(n\xi_n+\hat{\xi}_n\right) = 0. \quad (4.67)$$

So the relation between UV and IR renormalon residues of Eq. (4.38), which guarantees finiteness at $Q^2=\Lambda^2$, ensures that the characteristic function $\omega_{\mathcal{D}}(t)$ is continuous at $t=1$. Therefore, finiteness of $\mathcal{D}_{\rm PT}^{(L)}(\Lambda^2)$ would appear to stem from continuity of $\omega_{\mathcal{D}}(t=1)$.

4.4.1 Continuity of the Adler D-function and its derivatives at the Landau pole

For the first derivative of $\omega_{\mathcal{D}}(t)$, at $t=1$ one finds the discontinuity,

$$\left.\frac{d\omega_{\mathcal{D}}^{\rm IR}}{dt}\right|_{t=1} - \left.\frac{d\omega_{\mathcal{D}}^{\rm UV}}{dt}\right|_{t=1} = b\sum_{n=1}^{\infty} z_n^2[A_1(n)+B_1(n)] - \frac{b^2}{4}\sum_{n=1}^{\infty} z_n^2[A_0(n)+B_0(n)] + \frac{b^2}{2}\sum_{n=1}^{\infty} z_n^3[A_1(n)-B_1(n)].$$

(4.68)

Equation (4.39), which ensures that $\mathcal{D}_{\rm PT}^{(L)}(\Lambda^2)$ is finite, means that the first line of this expression vanishes. Furthermore, the second line also vanishes, ensuring continuity of the first derivative of $\omega_{\mathcal{D}}(t)$. This also guarantees that the $\mathcal{D}_{\rm PT}^{(L)'}(\Lambda^2)$ is finite (the prime denoting the derivative $d/d\ln Q$). Indeed, the required relation corresponding to the vanishing of the coefficient of the potentially divergent $\ln a$ term in $\mathcal{D}_{\rm PT}^{(L)'}(\Lambda^2)$ is,

$$\sum_{n=1}^{\infty}[2z_n^3(A_1(n)-B_1(n))-z_n^2(A_0(n)+B_0(n)] \;=\; 0. \tag{4.69}$$

So finiteness of the first derivative of $\mathcal{D}_{\rm PT}^{(L)}(Q^2)$ at $Q=\Lambda$, corresponds to continuity of the first derivative of $\omega(t)$ at $t=1$. Furthermore, Eq. (4.69) written in terms of ξ_n and $\hat{\xi}_n$ is simply Eq. (4.67), with an extra factor of -2. Consequently, the continuity of $\omega_{\mathcal{D}}(t)$ and its first derivative (corresponding to finiteness of $\mathcal{D}_{\rm PT}^{(L)}(\Lambda^2)$ and $\mathcal{D}_{\rm PT}^{(L)'}(\Lambda^2)$) both stem from a single relation, Eq. (4.67).

The second and third derivatives of $\omega_{\mathcal{D}}(t)$ are also continuous at $t=1$, and their discontinuities involve additional new structures built from combinations of the $A_{0,1}$ and $B_{0,1}$. Furthermore, to ensure finiteness of $\mathcal{D}_{\rm PT}^{(L)''}(Q^2)$ at $Q^2=\Lambda^2$, one requires the relation,

$$\sum_{n=1}^{\infty}[3z_n^4(A_1(n)+B_1(n))-2z_n^3(A_0(n)-B_0(n))] \;=\; 0, \tag{4.70}$$

and for finiteness of $\mathcal{D}_{\rm PT}^{(L)'''}(Q^2)$ at $Q^2=\Lambda^2$, one requires the relation,

$$\sum_{n=1}^{\infty}[4z_n^5(A_1(n)-B_1(n))-3z_n^4(A_0(n)+B_0(n))] \;=\; 0, \tag{4.71}$$

and these are both satisfied. Equations (4.70) and (4.71) are also the relations which ensure that the second and third derivatives of $\omega_{\mathcal{D}}(t)$ are continuous at $t=1$. Furthermore, they can both be derived from the following relation,

$$\sum_{n=1}^{\infty}\left(n^3\xi_n+3n^2\hat{\xi}_n\right) \;=\; 0. \tag{4.72}$$

We conclude that finiteness of $\mathcal{D}_{\rm PT}^{(L)''}(Q^2)$ and $\mathcal{D}_{\rm PT}^{(L)'''}(Q^2)$ is ensured by continuity of the second and third derivatives of $\omega_{\mathcal{D}}(t)$ at $t=1$ and that this in turn stems from the

relation in Eq. (4.72).

The fourth and higher derivatives of $\omega_{\mathcal{D}}(t)$ are discontinuous at $t = 1$, as noted in Ref. [130].

4.5 The skeleton expansion and non-perturbative effects

In this section we wish to consider in more detail, the compensation of ambiguities between renormalons and the OPE. The regular OPE is a sum over the contributions of condensates with different mass dimensions. As discussed in section 3.7, in the case of the Adler D-function, the dimension-four gluon condensate is the leading contribution,

$$G_0(a(Q^2)) = \frac{1}{Q^4}\langle 0|FF|0\rangle C_{GG}(a(Q^2)), \tag{4.73}$$

where $C_{GG}(a(Q^2))$ is the appropriate Wilson coefficient. In general, the n^{th} term in the OPE expansion of Eq. (4.3) will have the coefficient,

$$\mathcal{C}_n(a(Q^2)) = C_n[a(Q^2)]^{\delta_n}(1 + O(a)). \tag{4.74}$$

The exponent δ_n corresponds to the anomalous dimension of the condensate operator concerned. As was stated in section 3.7, these non-perturbative corrections cannot be unambiguously defined and this leads to C_n having an imaginary part so that $C_n = C_n^{(R)} \pm iC_n^{(I)}$. This mirrors the ambiguity in the perturbative part. Hence the nth term in the expansion in Eq. (3.103) will have an ambiguous imaginary part of the form,

$$\begin{aligned} \text{Im}[\mathcal{D}_{\text{NP}}] &= \pm iC_n^{(I)}[a(Q^2)]^{\delta_n}\left(\frac{\Lambda^2}{Q^2}\right)^n(1 + \mathcal{O}(a)) \\ &= \pm iC_n^{(I)}[a(Q^2)]^{\delta_n}\text{e}^{-2n/ba}(1 + \mathcal{O}(a)). \end{aligned} \tag{4.75}$$

Here, to obtain the second line we have used Eq. (1.113). If one considers an IR_n renormalon singularity in the Borel plane with the generic form $K_n/(1 - z/z_n)^{\gamma_n}$, then one finds an ambiguous imaginary part arising of the form,

$$\text{Im}[\mathcal{D}_{\text{PT}}] = \pm K_n\frac{\pi z_n^{\gamma_n}}{\Gamma(\gamma_n)}e^{-z_n/a(Q^2)}a^{1-\gamma_n}[1 + O(a)]. \tag{4.76}$$

Here, the $\pm$ ambiguity comes from routeing the contour above or below the real z-axis in the Borel plane (see Eq. (3.102)). This is structurally the same as the ambiguous OPE term in Eq. (4.74), and if we make the identification $C_n^{(I)} = K_n \pi z_n^{\gamma_n}/\Gamma(\gamma_n)$ and $\delta_n = 1 - \gamma_n$, then the PT Borel and NP OPE ambiguities can cancel against each other [147]. Taking a PV of the Borel integral corresponds to averaging over the $\pm$ possibilities.

For $Q^2 < \Lambda^2$ the modified expansion of Eq. (4.16) will have an n^{th} coefficient of the form,

$$\tilde{C}_n(a(Q^2)) = \tilde{C}_n[a(Q^2)]^{\tilde{\delta}_n}(1 + O(a)). \tag{4.77}$$

In this case the exponent $\tilde{\delta}_n$ is related to the anomalous dimension of four-fermion operators with mass dimension-6, which are associated with UV renormalons [160]. IR divergences associated with these render the imaginary part ambiguous, and $\tilde{C}_n = \tilde{C}_n^{(R)} \pm i\tilde{C}_n^{(I)}$. Thus the n^{th} term in the modified expansion will have an ambiguous imaginary part of the form,

$$\begin{aligned} \text{Im}[\mathcal{D}_{\text{NP}}] &= \pm i\tilde{C}_n^{(I)}[a(Q^2)]^{\tilde{\delta}_n}\left(\frac{Q^2}{\Lambda^2}\right)^n (1 + \mathcal{O}(a)) \\ &= \pm i\tilde{C}_n^{(I)}[a(Q^2)]^{\tilde{\delta}_n} \mathrm{e}^{2n/ba}(1 + \mathcal{O}(a)). \end{aligned} \tag{4.78}$$

The modified Borel representation of Eq. (4.13) has ambiguities arising from UV renormalons. Assuming that the UV_n singularity is of the form $\tilde{K}_n/(1 + z/z_n)^{\tilde{\gamma}_n}$ one finds,

$$\text{Im}[\mathcal{D}_{\text{PT}}] = \pm \tilde{K}_n \frac{\pi z_n^{\tilde{\gamma}_n}}{\Gamma(\tilde{\gamma}_n)} e^{z_n/a(Q^2)} a^{1-\tilde{\gamma}_n}[1 + O(a)]. \tag{4.79}$$

This is structurally the same as the ambiguity in the modified NP expansion coefficient in Eq. (4.77), and so if $\tilde{C}_n^{(I)} = \tilde{K}_n \pi z_n^{\tilde{\gamma}_n}/\Gamma(\tilde{\gamma}_n)$ and $\tilde{\delta}_n = 1 - \tilde{\gamma}_n$, then the ambiguities can be cancelled.

In the one-chain (leading-b) approximation the renormalons are single or double poles (corresponding to $\gamma_n = 1$ or $\gamma_n = 2$). Consequently the ambiguous imaginary parts in Eqs. (4.76) and (4.79) contain factors of $a^{1-\gamma}$ which are 1 or $1/a$, respectively. For the Adler D-function, $\text{Im}[\mathcal{D}_{\text{PT}}^{(L)}]$ is obtained by making the change $\text{Ei}(x) \to \text{Ei}(x) \pm i\pi$

in the second term in Eq. (4.31) for $Q^2 > \Lambda^2$, and in the first term for $Q^2 < \Lambda^2$. To ensure continuity of the imaginary part at $Q^2 = \Lambda^2$, one needs to choose the sign of $i\pi$ oppositely in the two regions. We then find, for $Q^2 > \Lambda^2$,

$$\begin{aligned}\mathrm{Im}[\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2)] &= \mp i\pi \left[\sum_{n=1}^{\infty} B_1(n+1) z_{n+1}^2 \left(\frac{\Lambda^2}{Q^2}\right)^{(n+1)} \right. \\ &\left. - \frac{1}{a(Q^2)} \sum_{n=2}^{\infty} z_{n+1}^2 [B_0(n+1) + z_{n+1} B_1(n+1)] \left(\frac{\Lambda^2}{Q^2}\right)^{n+1}\right].\end{aligned} \tag{4.80}$$

Correspondingly, for $Q^2 < \Lambda^2$ one finds,

$$\begin{aligned}\mathrm{Im}[\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2)] &= \pm i\pi \left[\sum_{n=2}^{\infty} A_1(n-1) z_{n-1}^2 \left(\frac{Q^2}{\Lambda^2}\right)^{n-1} \right. \\ &\left. - \frac{1}{a(Q^2)} \sum_{n=2}^{\infty} z_{n-1}^2 [A_0(n-1) - z_{n-1} A_1(n-1)] \left(\frac{Q^2}{\Lambda^2}\right)^{n-1}\right].\end{aligned} \tag{4.81}$$

Alternatively, we can obtain the same results by directly evaluating the residues along the integration contours in Eqs. (4.10) and (4.13) using the Borel transform of Eq. (4.25).

Comparing these expressions with Eqs. (4.64) and (4.65) one then finds that the imaginary part may be written directly in terms of the characteristic function $\omega_{\mathcal{D}}(t)$,

$$\begin{aligned}\mathrm{Im}[\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2)] &= \pm \frac{2\pi}{b} \frac{\Lambda^2}{Q^2} \omega^{\mathrm{IR}}_{\mathcal{D}}\left(\frac{\Lambda^2}{Q^2}\right) \qquad (Q^2 > \Lambda^2), \\ \mathrm{Im}[\mathcal{D}^{(L)}_{\mathrm{PT}}(Q^2)] &= \pm \frac{2\pi}{b} \frac{\Lambda^2}{Q^2} \omega^{\mathrm{UV}}_{\mathcal{D}}\left(\frac{\Lambda^2}{Q^2}\right) \qquad (Q^2 < \Lambda^2).\end{aligned} \tag{4.82}$$

Continuity of these expressions at $Q^2 = \Lambda^2$ then follows from continuity of $\omega(t)$ at $t = 1$. By requiring that these ambiguities are matched and cancelled by their OPE and modified expansion counterparts, we can relate the ambiguities in $\mathcal{D}^{(L)}_{\mathrm{NP}}(Q^2)$ to $\omega_{\mathcal{D}}(t)$ in both the UV and IR regions.

The $C_n^{(R)}$ and $\tilde{C}_n^{(R)}$ coefficients of the OPE and the modified NP expansion are in principle independent of the imaginary part. However, continuity at $Q^2 = \Lambda^2$ is dependent

upon relations between the $A_{0,1}$ and $B_{0,1}$ residues, such as Eqs. (4.38) and (4.40), together with the more complicated structures of Eqs. (4.69) - (4.71), which are required for finiteness of the Q^2 derivatives. Although not strictly necessary for continuity, this continuity follows naturally if we write,

$$\mathcal{D}^{(L)}_{\rm NP}(Q^2) \;=\; \left(\kappa \pm \frac{2\pi i}{b}\right) \int_0^{\Lambda^2/Q^2} dt \, \left(\omega_{\mathcal{D}}(t) + t\frac{d\omega_{\mathcal{D}}(t)}{dt}\right). \tag{4.83}$$

Here, κ is an undetermined overall real, non-perturbative factor and the t integration reproduces the expressions of Eq. (4.82) in the two Q^2 regions. If the perturbative component is PV regulated one averages over the $\pm$ possibilities. Combining Eq. (4.83) with Eq. (4.47) for $\mathcal{D}^{(L)}_{\rm PT}(Q^2)$ one can then write down a result for $\mathcal{D}^{(L)}(Q^2)$ for all values of Q^2,

$$\mathcal{D}^{(L)}(Q^2) \;=\; \int_0^{\infty} dt \, \left[\omega_{\mathcal{D}}(t)a(tQ^2) + \kappa\left(\omega_{\mathcal{D}}(t) + t\frac{d\omega_{\mathcal{D}}(t)}{dt}\right)\theta(\Lambda^2 - tQ^2)\right]. \tag{4.84}$$

The Q^2 evolution is then fixed by the non-perturbative constant κ, and by Λ. The IR limit is $\mathcal{D}^{(L)}(0) = 0$; we have already noted that $\mathcal{D}^{(L)}_{\rm PT}(0) = 0$, however the NP component also freezes to zero, since on integrating the second term, one finds an IR limit of $\omega^{\rm IR}_{\mathcal{D}}(1) - \omega^{\rm UV}_{\mathcal{D}}(1) = 0$, from continuity of the characteristic function at $t = 1$. The same expression holds for the other Euclidean observables $\mathcal{K}^{(L)}_{\rm PT}(Q^2)$ and $\mathcal{U}^{(L)}_{\rm PT}(Q^2)$ on replacing $\omega_{\mathcal{D}}(t)$ by $\omega_{\mathcal{K}}(t)$ and $\omega_{\mathcal{U}}(t)$, respectively.[3] From Eq. (4.84) we can determine that,

$$\mathcal{D}^{(L)}(\Lambda^2) \;=\; \frac{0.679938\ldots}{b} + \frac{4}{3}\left(\frac{7}{2} - \frac{\pi^2}{3}\right). \tag{4.85}$$

We plot in Fig. 4.4 the overall result for $\mathcal{D}^{(L)}(Q^2)$, $\mathcal{K}^{(L)}(Q^2)$ and $\mathcal{U}^{(L)}(Q^2)$ for the choices $\kappa = 0$, $\kappa = 1$ and $\kappa = -1$. For the DIS sum rules, $\omega_{\mathcal{K}}(t)$, $\omega_{\mathcal{U}}(t)$ and their first derivatives are continuous at $t = 1$. In the case of $\mathcal{U}^{(L)}_{\rm NP}(Q^2)$ there are a total of three non-perturbative terms, and hence the continuity of the characteristic function and its first derivative fixes the form of the function up to an overall constant factor. Thus Eq. (4.84) does indeed hold for $\mathcal{U}^{(L)}(Q^2)$ without conjecturing the form of Eq. (4.83).

[3]The characteristic functions for $\mathcal{K}(a)$ and $\mathcal{U}(a)$ can be obtained by taking the Borel transforms of Eqs. (4.27) and (4.28), substituting these into Eq. (4.10) and then reversing the changes of variables used to obtain Eq. (4.59).

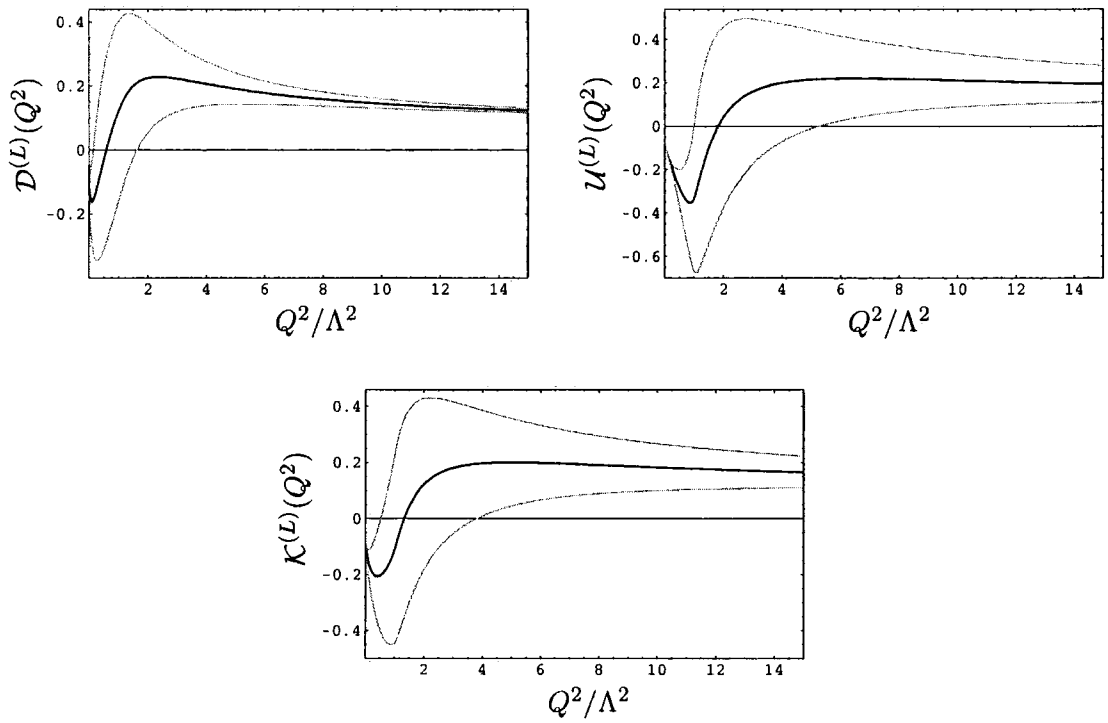


Figure 4.4: Here we plot the combined perturbative plus non-perturbative corrections to the observables, which are evaluated using Eq. (4.84). In each of the plots we adopt three different values of the non-perturbative parameter κ. The grey curves correspond to $\kappa = -1$ and $\kappa = 1$ and the bold curve corresponds to $\kappa = 0$ (equivalent to the expressions in Eqs. (4.31) - (4.33)).

4.6 Infrared freezing behaviour of the $R_{e^+e^-}$ ratio

We turn in this section to a consideration of the freezing behaviour of the Minkowskian quantity $R_{e^+e^-}$ which was discussed in Ref. [48]. This treatment was criticized in Ref. [161], which argued that in fact there is an unphysical divergence in the IR limit. We wish to address these criticisms.

$R_{e^+e^-}(s)$ will be defined by Eq. (4.21) with the perturbative corrections $\mathcal{D}(Q^2)$ replaced by $\mathcal{R}(s)$, where $\sqrt{s}$ is the e^+e^- c.m. energy. $\mathcal{R}(s)$ is related to $\mathcal{D}(-s)$ by analytical continuation from Euclidean to Minkowskian space. One may write the dispersion relation,

$$\mathcal{R}(s) = \frac{1}{2\pi i}\int_{-s-i\epsilon}^{-s+i\epsilon} dt \frac{\mathcal{D}(t)}{t}. \tag{4.86}$$

If $\mathcal{D}(t)$ is represented by a Borel representation as in Eq. (4.10) then one arrives at,

$$\mathcal{R}^{(L)}_{\rm PT}(s) = \int_0^\infty dz\, e^{-z/a(s)} \frac{\sin(\pi bz/2)}{\pi bz/2} B[\mathcal{D}^{(L)}_{\rm PT}](z). \tag{4.87}$$

There is now an extra oscillatory factor of $\sin(\pi bz/2)/(\pi bz/2)$ arising from the analytical continuation. In consequence, each *individual* IR or UV renormalon contribution at $Q^2 = \Lambda^2$ will be finite, and the cancellation of Eq. (4.39) is not required. One can also analytically continue the one-chain skeleton expansion result for $\mathcal{D}^{(L)}_{\rm PT}(Q^2)$ to obtain,[4]

$$\mathcal{R}^{(L)}_{\rm PT}(s) = \frac{2}{\pi b}\int_0^\infty dt\, \omega_{\mathcal{D}}(t) \arctan\left(\frac{\pi b a(ts)}{2}\right). \tag{4.88}$$

Here, the principal branch of arctan is assumed so that it lies in the interval $[-\pi/2, +\pi/2]$, and $\arctan(0) = 0$. This form is equivalent to the Borel representation of Eq. (4.87) for $s > \Lambda^2$, and to the modified Borel representation for $s < \Lambda^2$. Notice that the choice of principal branch is crucial if the PV Borel sum is to be continuous at $s = \Lambda^2$. The result freezes to the IR limit $\mathcal{R}^{(L)}_{\rm PT}(0) = 0$, since $\arctan(0) = 0$ on the principal branch. This freezing limit differs from that found in the 'analytic perturbation theory' approach [162–164], where freezing to an IR limit of $2/b$ occurs. This freezing limit was also erroneously claimed in Ref. [48], but unfortunately leads to the PV Borel sum being discontinuous. In Ref. [161], unphysical singularities in the region $-\Lambda^2 < s < 0$ lead to extra terms and they find,

$$\begin{aligned}\mathcal{R}^{(L)}_{\rm PT}(s) &= \frac{2}{\pi b}\int_0^\infty dt\, \omega_{\mathcal{D}}(t) \arctan\left(\frac{\pi b a(ts)}{2}\right) \\ &\quad + \frac{2}{b}\int_0^{\Lambda^2/s} dt\, \omega_{\mathcal{D}}(t) + \frac{2}{b}\int_{-\Lambda^2/s}^0 dt\, \omega^{\rm IR}_{\mathcal{D}}(t).\end{aligned} \tag{4.89}$$

The extra terms may be treated as contributions to $\mathcal{R}^{(L)}_{\rm NP}(Q^2)$. The final term leads to an IR divergence as $s \to 0$, and has an expansion of the same form as the OPE. Note however, that the Minkowskian OPE for $\mathcal{R}(Q^2)$ is pathological and contains delta-functions, $\delta(s)$, and their derivatives [165]. It is only when a smearing procedure in Q^2 is used [166] that it makes sense. In contrast, for Euclidean quantities the regular OPE is potentially well-defined, and no smearing is required.

We will now consider the evaluation of the PV Borel integral for $\mathcal{R}^{(L)}_{\rm PT}$, and correct the

[4] Note that for $R_{e^+e^-}$, we cannot obtain a skeleton expansion representation of the form of Eq. (4.6). The appropriate characteristic function, $\omega_{\mathcal{R}}(t)$ does not exist.

erroneous statements made in Ref. [48] and noted above. This can be expressed in terms of generalized exponential integral functions $\mathrm{Ei}(n,w)$, defined for $\mathrm{Re}\, w > 0$ by

$$\mathrm{Ei}(n,w) \;=\; \int_1^\infty dt \frac{e^{-wt}}{t^n}. \tag{4.90}$$

One also needs the integral,

$$\int_0^\infty dz\, e^{-z/a} \frac{\sin(\pi b z/2)}{z} \;=\; \arctan\left(\frac{\pi b a}{2}\right). \tag{4.91}$$

Here, the principal branch of the arctan is again assumed. Care needs to be taken when $\mathrm{Re}\, w < 0$. With the standard continuation one arrives at a function which is analytic everywhere in the cut complex w-plane, except at $w = 0$, and with a branch cut running along the negative real semi-axis. Explicitly [167],

$$\mathrm{Ei}(n,w) \;=\; \frac{(-w)^{n-1}}{(n-1)!}\left[-\ln w - \gamma_E + \sum_{m=1}^{n-1}\frac{1}{m}\right] - \sum_{\substack{m=0 \\ m\neq n-1}} \frac{(-w)^m}{(m-n+1)m!}. \tag{4.92}$$

The $\ln w$ term contributes to the branch cut along the negative real semi-axis. To obtain the PV of the Borel integral one needs to compensate for the discontinuity across the branch cut, and make the replacement $\mathrm{Ei}(n,w) \to \mathrm{Ei}(n,w) + i\pi\, \mathrm{sign}(\mathrm{Im}\, w)$. One then finds that the IR renormalon contributions to $\mathcal{R}^{(L)}_{\rm PT}(s)$ can be written in terms of the function,

$$\begin{aligned}\phi_-(p,q) \;&\equiv\; e^{-z_q/a(s)}(-1)^q \mathrm{Im}[\mathrm{Ei}(p, -z_q/a(s) - i\pi b z_q/2)] \\ &-\; \frac{e^{-z_q/a(s)}(-1)^q z_q^{p-1}}{(p-1)!}\pi \mathrm{Re}[((z_q/a(s)) + i\pi b/2)^{p-1}]\, \theta(s-\Lambda^2). \end{aligned}\tag{4.93}$$

The UV renormalon contributions can be written in terms of the function,

$$\begin{aligned}\phi_+(p,q) \;&\equiv\; e^{z_q/a(s)}(-1)^q \mathrm{Im}[\mathrm{Ei}(p, z_q/a(s) + i\pi b z_q/2)]. \\ &+\; \frac{e^{z_q/a(s)}(-1)^q z_q^{p-1}}{(p-1)!}\pi \mathrm{Re}[((-z_q/a(s)) - i\pi b/2)^{p-1}]\, \theta(\Lambda^2 - s). \end{aligned}\tag{4.94}$$

The PV regulated $\mathcal{R}^{(L)}_{\rm PT}(s)$ is then given for *all* values of s by,

$$\begin{aligned}\mathcal{R}^{(L)}_{\rm PT}(s) \;&=\; \mathcal{R}^{(L)}_{\rm PT}(s)|_{\rm UV} + \mathcal{R}^{(L)}_{\rm PT}(s)|_{\rm IR} \\ &=\; \frac{2}{\pi b}\arctan\left(\frac{\pi b a(s)}{2}\right)\end{aligned}$$

$$+\frac{2}{\pi b}\sum_{j=1}^{\infty}(A_0(j)\phi_+(1,j)+(A_0(j)-A_1(j)z_j)\phi_+(2,j))$$

$$+\frac{2B_0(2)}{\pi b}\phi_-(1,2)+\frac{2}{\pi b}\sum_{j=3}^{\infty}(B_0(j)\phi_-(1,j)+(B_0(j)+B_1(j)z_j)\phi_-(2,j))\,. \tag{4.95}$$

Note that the presence of the θ-functions is crucial in Eqs. (4.93) and (4.94). The terms they multiply are the extra contributions necessary to obtain the PV when Re $w < 0$. For $s > \Lambda^2$ the second contribution is required for the IR renormalon contribution, but for $s < \Lambda^2$ it must be switched off, otherwise the Borel integral will not be correctly evaluated. With $a(s) < 0$ for $s < \Lambda^2$, Re $w < 0$ occurs for the UV renormalon contributions and the extra term must be switched on to obtain a PV regulation of the UV component. Leaving out the θ-function in Eq. (4.93) would cause an unphysical divergence in the IR, and leaving it out in Eq. (4.94) would cause asymptotic freedom to fail in the UV. If the PV is correctly evaluated with arctan remaining on the principal branch for $s^2 < \Lambda^2$, then one obtains $\mathcal{R}^{(L)}_{\rm PT}(0) = (2/\pi b)\arctan(0) = 0$. Notice that at first sight the PV result appears to be discontinuous at $s = \Lambda^2$, as the θ-function contributions switch over. However the discontinuity is given by the $\phi_\pm(1,j)$ terms, and one finds, upon summing them, a discontinuity,

$$\frac{2}{b}\sum_{j=1}^{\infty}[B_0(j)(-1)^j + A_0(j)(-1)^j] \;=\; \frac{2}{b}B_0(2) \;=\; \frac{2}{b}. \tag{4.96}$$

Here, the relation of Eq. (4.40) ensures pairwise cancellations of terms, and $B_0(2) = 1$ is left over. If we remain on the principal branch, however, the arctan term has an equal discontinuity of opposite sign, since $(2/\pi b)\arctan(+\infty) = 1/b$, whereas $(2/\pi b)\arctan(-\infty) = -1/b$, and overall there is continuity at $s = \Lambda^2$. In Ref. [48] it was wrongly claimed that the PV result is discontinuous at $s = \Lambda^2$, and instead it was suggested that a regulation in which one throws away the second terms in Eqs. (4.93) and (4.94) should be used. These terms are of the form $(\Lambda^2/s)^q$, and $(s/\Lambda^2)^q$, respectively, and so they can simply be absorbed into the regular OPE and its modified form.

We finally discuss the ambiguous part of $\mathcal{R}^{(L)}_{\rm PT}(s)$, given by $\mathrm{Im}[\mathcal{R}^{(L)}_{\rm PT}(s)]$. This may be evaluated straightforwardly as,

$$\mathrm{Im}[\mathcal{R}^{(L)}_{\rm PT}(s)] \;=\; \pm i\pi\sum_{n=1}^{\infty}[B_0(n)+B_1(n)z_n]z_n(-1)^n\left(\frac{\Lambda^2}{s}\right)^n \qquad (s>\Lambda^2),$$

$$\mathrm{Im}[\mathcal{R}^{(l)}_{\mathrm{PT}}(s)] \;=\; \mp i\pi \sum_{n=1}^{\infty}[A_0(n) - A_1(n)z_n]z_n(-1)^n\left(\frac{\Lambda^2}{s}\right)^{-n} \qquad (s < \Lambda^2). \tag{4.97}$$

If one defines $\omega_{\mathcal{D}}(t) \equiv \omega^{(1)}_{\mathcal{D}}(t) + \ln t\, \omega^{(2)}_{\mathcal{D}}(t)$ (the split being into the single and double pole renormalon contributions), then comparing with Eqs. (4.64) and (4.65) one finds,

$$\begin{aligned}
\mathrm{Im}[\mathcal{R}^{(L)}_{\mathrm{PT}}(s)] \;&=\; \pm\frac{2\pi}{b}\frac{\Lambda^2}{s}\omega^{(2),\,\mathrm{IR}}_{\mathcal{D}}\left(\frac{-\Lambda^2}{s}\right) \qquad (s > \Lambda^2), \\
\mathrm{Im}[\mathcal{R}^{(L)}_{\mathrm{PT}}(s)] \;&=\; \pm\frac{2\pi}{b}\frac{\Lambda^2}{s}\omega^{(2),\,\mathrm{UV}}_{\mathcal{D}}\left(\frac{-\Lambda^2}{s}\right) \qquad (s < \Lambda^2).
\end{aligned} \tag{4.98}$$

Notice that only the double poles contribute since the $\sin(\pi bz/2)/(\pi bz/2)$ analytical continuation term in Eq. (4.87) contains zeros at $z = \pm z_n$ which nullify the single pole contributions. Whilst the characteristic function $\omega_{\mathcal{D}}(t)$ is continuous at $t = 1$, the $\omega^{(2)}_{\mathcal{D}}(t)$ function is *discontinuous* at $t = -1$. The magnitude of the discontinuity is $\pm 2/b$ and arises from the same sum in Eq. (4.96), which gave an apparent discontinuity in the PV regulated $\mathcal{R}^{(L)}_{\mathrm{PT}}(s)$ component, although in the PT case this is cancelled by the arctan term. Thus, defined in this way, $\mathrm{Im}[\mathcal{R}^{(L)}_{\mathrm{PT}}(s)]$ is discontinuous at $s = \Lambda^2$. It would seem that the proper way to proceed is to use the dispersion relation of Eq. (4.86) to analytically continue the expression for $\mathcal{D}^{(L)}(Q^2)$ arrived at in Eq. (4.84) into the Minkowskian region. Unfortunately, the one-chain skeleton expansion form for $\mathcal{D}(Q^2)$ is difficult to analytically continue in a consistent fashion, which was a key motivation for the alternative inverse Mellin representation introduced in Ref. [168]. Clearly the issue of Minkowskian freezing is more subtle and requires further investigation.

4.7 Summary and Conclusions

We have shown in this chapter that the leading-b, all-orders resummed form of Euclidean observables are finite at the Landau pole, despite the singular behaviour of the coupling at this point. In addition, we found that perturbative corrections to the observables remain finite below the Landau pole, and freeze to zero. Previous studies of potentially finite behaviour of these observables at or below Λ^2, such as the 'analytic perturbation theory' approach [162–164], have focused on the use of a modified or 'analytic' coupling, which is itself finite all the way to $Q^2 = 0$ and this automatically results in finite IR behaviour. However, we have shown that this is not necessary. When one

sums the perturbative expansion to all-orders, the logarithmic divergences in the coupling at Λ^2 cancel, leaving the full result finite. Furthermore, we have shown that this is due to a curious relation between the IR and UV behaviour of QCD – cancellation between UV and IR renormalons. Whilst we might have expected the full form of the perturbative expansion[5] to have finite IR properties, it is surprising that we can obtain such a result in the large-b limit, and using only the one-loop form of the coupling. The structure of IR$\leftrightarrow$UV relations between the residues of renormalon singularities in the Borel plane appears to be much more extensive than previously thought. This finite behaviour also confirms the leading-b expansion as being of special significance in QCD.

When we study the skeleton expansion representation of these observables, we can relate these structures to the continuity of the characteristic function – a function which we can interpret physically as the distribution of momentum flowing through the renormalon chains in Figs. 3.3 and 4.1, and the form of which is dictated by the nature of the skeleton diagram itself.

We found that before freezing to zero, the perturbative corrections change sign in the vicinity of the Landau pole. However, when added to the parton model result, the observable remain positive for all Q^2. We also demonstrated that the skeleton expansion language provided a natural link between the standard Borel representation (valid in the $Q^2 > \Lambda^2$ region) and a modified Borel representation (applicable to the $Q^2 < \Lambda^2$ region). In doing so, we established a dictionary between the residues of renormalon singularities and the series expansion coefficients of the characteristic function. We showed that for the modified representation, the singularity structure of the Borel plane is inverted, with UV renormalons now being the source of ambiguities. This further highlights the strong relation between UV and IR physics embedded in the renormalon description of QCD.

We discussed the ambiguous imaginary part of the renormalon contribution to the perturbative expansion, and how this differs between the UV and IR regions. We then developed a model for non-perturbative power corrections, based on the compensation of perturbative and non-perturbative ambiguities. This was only made possible by the newly discovered attractive IR properties of the renormalon component noted

[5]By this we mean the form of Eq. (4.1) including *all* higher-order coefficients in their exact form and the full all-orders coupling.

above. The result of this is an expression for the combined perturbative and non-perturbative components of the observable, written in terms of the QCD scale parameter Λ and a single undetermined non-perturbative parameter, κ.

All of the above properties and results hold in general for Euclidean observables for which a one-chain result of the form of Eq. (4.6) can be written down.

We also briefly considered the IR properties of the $R_{e^+e^-}$ ratio, for which a characteristic function, $\omega_{\mathcal{R}}(t)$, in the representation of Eq. (4.6) does not exist. We concluded that the subject of IR freezing in the case of Minkowskian observables requires further study.

Having developed these theoretical predictions, an obvious next step is to apply them to a phenomenological analysis of the observables concerned. Experimental data exists for the GLS and polarized Bjorken sum rules and data for the Adler D-function can be obtained indirectly via analytical continuation of data for the $R_{e^+e^-}$ ratio. A study of this kind (in the case of the sum rules) is the subject of the next chapter. Also, it is interesting to note that the characteristic function of the Adler D-function is continuous only up to its third derivative. This would presumably cause problems with the continuity and finiteness of the separate perturbative and non-perturbative components of the fourth derivative of the Adler D-function, when written in terms of Eq. (4.84). It is unclear whether we can make any conclusions about the physical implications of this and this subject requires further study.

One might expect that the IR properties of the one-chain result also hold for the two-chain result. However, the relation between the two-chain form of Eq. (4.6) and the sub-leading-N_f and -b coefficients $d_n^{[n-1]}$ and $d_n^{(n-1)}$ is slightly more complicated. Whereas the one-chain term corresponds exactly to the leading-N_f coefficient, $d_n^{[n]}$, only part of the sub-leading-N_f coefficient, $d_n^{[n-1]}$ is built from two-chain contributions. This makes it difficult to draw any conclusions on general grounds about the IR properties of the two-chain term in the skeleton expansion. Clearly, the nature of the two-chain term also warrants further study.

Chapter 5

Comparison of NNLO and all-orders estimates of corrections to QCD sum rules

In this chapter we build upon the progress made in chapter 4 by using the results obtained there to perform a phenomenological analysis of QCD sum rules. We apply CORGI perturbation theory to the all-orders leading-b resummation of the GLS and Bjorken sum rules, and compare resulting predictions with experimental data. We also test the proposed model for non-perturbative power corrections and perform fits to obtain an optimal value of the non-perturbative parameter in each case.

5.1 Introduction

As we have seen, perturbation theory can only remain well-defined when supplemented by non-perturbative corrections in the form of the OPE. Even in the limit of asymptotic freedom, non-perturbative corrections are needed in order to account for the existence of ambiguous imaginary components of perturbative predictions. Consequently, for low values of Q^2, not only do the convergence properties of fixed order estimates suffer, but we also have to contend with a large theoretical error associated with the intrinsically ambiguous nature of perturbation theory (see Eqs. (3.28) and (3.102)). The inclusion of non-perturbative contributions should act to negate this error.

Furthermore, we can assess the likely accuracy of fixed-order predictions by comparing them with their all-orders counterparts. In the past, this has been done for quantities such as the $R_{e^+e^-}$ ratio at low values of the c.m. energy (where the QCD coupling is not so small), and for its tau-decay analogue R_τ, where again, at the scale of the τ mass, the coupling is relatively large [130, 143, 169–171]. DIS sum rules have also been studied in Refs. [172–179]. However, in matching to the exact fixed-order perturbative calculations (NLO and NNLO coefficients are available for $R_{e^+e^-}$, R_τ and DIS sum rules), one faces the problem that the all-orders leading-b result is only renormalization scheme (RS) invariant if the one-loop form of the coupling is used, whereas the exact result must involve the higher-loop coupling. The matched leading-b resummed result consequently depends on the assumed renormalization scale at which the matching takes place. To avoid this ambiguity it has been argued in the past that one should remove scale dependence by a resummation of scale logs to all-orders; employing the effective charge approach [41,42], or the CORGI approach [38,39]. The scheme invariant quantities associated with these approaches (given in Eqs. (1.158) - (1.159)) – like the standard perturbative coefficients – can be approximated at the leading-b level and, subsequently resummed [170, 171, 180].

The sum rules we will consider have only been measured for relatively small values of Q^2 [97, 102, 181–184]. Consequently, for the reasons stated above, non-perturbative corrections must be included in order to perform an accurate phenomenological comparison. In light of this, we use the non-perturbative model developed in the previous chapter to improve the accuracy of QCD predictions and to quantify the importance of non-perturbative effects for the individual sum rules.

The work presented in this chapter is based on the research carried out in Ref. [185] and the material is organized as follows:

We describe how the CORGI approach can be adapted to leading-b resummations of GLS, polarized and unpolarized Bjorken sum rules. We then compare the resultant predictions with the available data from experiment and assess the consistency between fixed-order and all-orders approaches. In so doing it is necessary to choose a method for evolving through quark mass thresholds and we describe the method we adopt for this analysis, and for the remainder of this thesis. We then apply the model for power corrections developed in the previous chapter to these observables and perform fitting

for the non-perturbative parameter κ. Finally we discuss what conclusions can be drawn from this work.

5.2 Fixed-order and all-orders predictions

The three physical quantities we are interested in are the GLS sum rule [81], the polarized Bjorken sum rule [90], and the unpolarized Bjorken sum rule [85]. These quantities were defined in section 2.7.1 and can be written as a parton model result plus perturbative corrections (denoted by calligraphic letters). The perturbative corrections take the form of expansions in powers of the strong coupling $a = \alpha_s/\pi$,

$$\mathcal{K}(a) = a(Q) + \sum_{n=1}^{\infty} k_n a^{n+1}(Q), \qquad \mathcal{U}(a) = a(Q) + \sum_{n=1}^{\infty} u_n a^{n+1}(Q), \quad (5.1)$$

with $\mathcal{K}(a)$ being relevant to the GLS and polarized Bjorken sum rules and $\mathcal{U}(a)$ to the unpolarized Bjorken sum rule. These expressions can be evaluated at fixed-order, and the coefficients k_n and u_n are known up to $n = 2$ (NNLO) [82–84, 86–88].

$$\mathcal{K}_{NNLO}(a) = a + k_1 a^2 + k_2 a^3, \qquad \mathcal{U}_{NNLO}(a) = a + u_1 a^2 + u_2 a^3. \quad (5.2)$$

The leading-b approximation (previously discussed in chapters 3 and 4) arises from rearranging the coefficients into an expansion in powers of N_f. The leading N_f term can be obtained exactly to all-orders and converting this into the leading-b term allows us to approximate the full coefficients at all orders. The result is a Borel integral which resums all leading-b terms $k_n^{(L)} \equiv k_n^{(n)} b^n$ and $u_n^{(L)} \equiv u_n^{(n)} b^n$. In the case of $\mathcal{K}(a)$, this has the form,[1]

$$B[\mathcal{K}^{(L)}](z) = \sum_{n=0}^{\infty} \frac{k_n^{(L)}}{n!} z^n \qquad \Longrightarrow \qquad \mathcal{K}_\infty(a) \simeq \int_0^\infty dz e^{-z/a} B[\mathcal{K}^{(L)}](z), \quad (5.3)$$

with a similar expression for $\mathcal{U}_\infty(a)$.

The Borel transforms of the above sum rules in the leading-b approximation can be

[1] In this chapter we have changed notation slightly. We redefine the all-orders resummations as $\mathcal{K}_\infty(a) \equiv \mathcal{K}_{\rm PT}^{(L)}(a)$ and $\mathcal{U}_\infty(a) \equiv \mathcal{U}_{\rm PT}^{(L)}(a)$. This emphasizes the number of terms in Eq. (5.1) included in these predictions (relative to the NLO and NNLO expressions) and also makes the labelling of the plots less cluttered.

calculated from the results in Refs. [135] and [159], and they are given in Eqs. (4.27) and (4.28). One can easily compute the PV regulated leading-b Borel sums [143] and these are given in Eqs. (4.32) and (4.33). Here, we denote them by, $\mathcal{K}_{\infty}(Q^2)$ and $\mathcal{U}_{\infty}(Q^2)$.

For these all-orders resummations to be RS (scale)-invariant we need to take $a(Q^2)$ as the one-loop coupling, given in Eq. (4.2). Also, Eqs. (4.32) and (4.33) are derived in the so-called V-scheme ($\overline{\rm MS}$ subtraction with renormalization scale $\mu^2 = e^{-5/3}Q^2$), and hence Λ in Eq. (4.2) will refer to that in the V-scheme defined by,

$$\Lambda_{\rm V} \;=\; {\rm e}^{5/6}\Lambda_{\overline{\rm MS}}. \tag{5.4}$$

As noted above, there is a problem if we wish to match these leading-b resummations with exact higher-order NLO and NNLO calculations for $\mathcal{K}(a)$ and $\mathcal{U}(a)$ which involve the higher-loop coupling. We shall avoid the matching ambiguity by employing the CORGI approach described in section 1.8.3 [39, 186]. The CORGI approach consists of resumming to all-orders the RG-predictable terms available to a given fixed-order of calculation (see section 1.8.3). The standard perturbative expansion is then replaced with the following expression,

$$\tilde{\mathcal{K}}(a_0) \;=\; a_0 + \sum_{n=2}^{\infty} X_n a_0^{n+1}, \tag{5.5}$$

which removes the renormalization scale dependence of Eq. (5.1). In this chapter, we will use a tilde to denote CORGI-ized results. Initially we shall deal with $\mathcal{K}(a)$ but the results are easily generalized to $\mathcal{U}(a)$.

The coefficients X_n are renormalization scheme invariant quantities, each of which can be derived from an N^{n}LO calculation of the coefficients of $\mathcal{K}(a)$, and of the beta function equation coefficients, c_n. They are given for $n = 2$ and 3 in Eqs. (1.158) - (1.159) and in the case of $\mathcal{K}(a)$, they have the form,

$$X_2 \;=\; k_2 - k_1^2 - ck_1 + c_2, \tag{5.6}$$

$$X_3 \;=\; k_3 - 3k_1k_2 + 2k_1^3 + \frac{ck_1^2}{2} - k_1c_2 + \frac{1}{2}c_3. \tag{5.7}$$

The coupling a_0 used in Eq. (5.5) has the same form as that given in Eq. (1.152) [35,36],

$$a_0(Q) = \frac{-1}{c\Big(1+W_{-1}\Big[-\frac{1}{e}\Big(\frac{Q^2}{\Lambda_{\mathcal{K}}^2}\Big)^{-b/2c}\Big]\Big)}, \tag{5.8}$$

but in terms of the scale $\Lambda_{\mathcal{K}}$. The relation between $\Lambda_{\mathcal{K}}$ and the $\overline{\rm MS}$ scale can be obtained through the expression for X_0, evaluated in $\overline{\rm MS}$ with $\mu = Q$,

$$X_0(Q) = b\ln\left(\frac{\mu}{\tilde{\Lambda}_{\overline{\rm MS}}}\right) - k_1^{\overline{\rm MS}} \tag{5.9}$$

$$= b\ln\frac{Q}{\Lambda_{\mathcal{K}}} \tag{5.10}$$

$$\implies \Lambda_{\mathcal{K}} = \Lambda_{\overline{\rm MS}}\left(\frac{2c}{b}\right)^{-c/b}\exp(k_1^{\overline{\rm MS}}/b)$$

$$\approx 2.439\Lambda_{\overline{\rm MS}}. \tag{5.11}$$

Here, $k_1^{\overline{\rm MS}}$ denotes the NLO perturbative coefficient in $\overline{\rm MS}$ with scale choice $\mu = Q$ and on the second line we have assumed $N_f = 4$. The exponential term in the above equation has the effect of resumming a set of RG predictable terms present in the full perturbative expansion, it also reproduces the $\mathcal{O}(a^2)$ term present in Eq. (5.1) but absent from Eq. (5.5).

Equation (5.5) can easily be adapted to provide a prediction for the unpolarized Bjorken sum rule of Eq. (4.24). We simply use the equivalent coefficients of $\mathcal{U}(a)$ in Eqs. (5.6) and (5.7) and the following coupling,

$$a_0(Q) = \frac{-1}{c\Big(1+W_{-1}\Big[-\frac{1}{e}\Big(\frac{Q^2}{\Lambda_{\mathcal{U}}^2}\Big)^{-b/2c}\Big]\Big)}, \tag{5.12}$$

where,

$$\Lambda_{\mathcal{U}} = \tilde{\Lambda}_{\overline{\rm MS}}\exp(u_1^{\overline{\rm MS}}/b)$$

$$\approx 2.901\Lambda_{\overline{\rm MS}}. \tag{5.13}$$

Again, we have assumed $N_f = 4$ in Eq. (5.13).

Equation (4.32) can be adapted so that it resums the leading-b components of Eq. (5.5)

and essentially becomes an all-orders CORGI result. We define the coupling a_v through,

$$\frac{1}{a_v} = \frac{1}{a_0} + k_1^{(1)} b, \tag{5.14}$$

where $k_1^{(1)}$ is defined through the k_n analogue of Eq. 4.4; it is calculated in the V-scheme[2] and has a value of 1/6. For $\mathcal{U}(a)$ we have $u_1^{(1)} = 1/2$. The coupling a_0 in Eq. (5.14) is that of Eq. (5.8). The all-orders CORGI result can now be obtained from Eq. (4.32) using Eq. (5.14),

$$\begin{aligned} \tilde{\mathcal{K}}_\infty(Q) &\equiv a_0 + \sum_{n=2}^{\infty} X_n^{(n)} b^n a_0^{n+1} \\ &= \mathcal{K}_\infty(a_v), \end{aligned} \tag{5.15}$$

with $\mathcal{K}_\infty(a)$ given by Eq. (4.32). $X_n^{(n)}$, in analogy with Eq. (3.70), is the leading-b part of X_n, for example:

$$X_2^{(2)} = k_2^{(2)} - \left(k_1^{(1)}\right)^2, \tag{5.16}$$

$$X_3^{(3)} = k_3^{(3)} - 3k_1^{(1)}k_2^{(2)} + 2\left(k_1^{(1)}\right)^3. \tag{5.17}$$

We can verify Eq. (5.15) by substituting for a_v using Eq. (5.14) and expanding in powers of a_0. This yields,

$$\begin{aligned} \tilde{\mathcal{K}}_\infty(Q) &\equiv a_v + \sum_{n=1}^{\infty} k_1^{(1)} b^n a_v^{n+1} \\ &= a_0 + a_0^3 b^2 \left(k_2^{(2)} - 2\left(k_1^{(1)}\right)^2\right) \\ &+ a_0^4 b^3 \left(k_3^{(3)} - 3k_1^{(1)}k_2^{(2)} + 2\left(k_1^{(1)}\right)^3\right) + \ldots, \end{aligned} \tag{5.18}$$

which is consistent with Eqs. (5.15) - (5.17).

Furthermore, we can also improve upon $\tilde{\mathcal{K}}_\infty$ by adding to Eq. (5.15) the known sub-leading-b part of X_2,

$$\tilde{\mathcal{K}}_{\infty+}(Q) = \tilde{\mathcal{K}}_\infty(Q) + (X_2 - X_2^{(2)} b^2) a_0^3(Q). \tag{5.19}$$

[2]To obtain the coefficients k_n in the V-scheme, or indeed in any scheme, we use the fact that X_n are RS invariant, in conjunction with Eqs. (5.9), (5.6) and (5.7).

Equation (5.19) now contains all the information we have about the perturbative coefficients for $\mathcal{K}(a)$ at all-orders. Analogous expressions also hold for $\tilde{\mathcal{U}}_{\infty}(Q)$ and $\tilde{\mathcal{U}}_{\infty+}(Q)$.

In chapter 4, we noted that the one chain result of Eq. (4.32) is finite at the Landau pole ($Q = \Lambda$), and remains so for values of Q below Λ. Remarkably, this finiteness at $Q = \Lambda$ holds when we use the 't Hooft coupling of Eq. (5.8). Furthermore, both of these results have the same values at their respective Landau poles, i.e.

$$\mathcal{K}(a)\Big|_{Q=\Lambda_v} = \mathcal{K}(a_0)\Big|_{Q=\Lambda_{\mathcal{K}}} = -\frac{8}{9}\ln 2. \tag{5.20}$$

Similar relations apply to $\mathcal{U}(a)$ and also to the Adler D-function [152]. This conclusion is altered slightly for the actual CORGI result of Eq. (5.15) because of the use of $a_v(Q)$. In this case the result remains finite at $Q = \Lambda_{\mathcal{K}}$, but has a different value to that in Eq. (5.20).

Finally, we can also consider NLO and NNLO CORGI, fixed-order approximations,

$$\tilde{\mathcal{K}}_{NLO}(a_0) = a_0(Q), \tag{5.21}$$

$$\tilde{\mathcal{K}}_{NNLO}(a_0) = a_0(Q) + X_2 a_0^3(Q). \tag{5.22}$$

By comparing these with the all-orders resummations, we can assess the reliability of fixed-order perturbative predictions.

5.3 Comparison of fixed-order and all-orders predictions

The data available for $K_{GLS}(Q)$ and $K_{uBj}(Q)$ span a range of energy which includes the bottom quark mass threshold. It is therefore necessary for us to choose a method for evolving the above expressions through this threshold and we adopt the approach detailed in Ref. [187]. At a particular energy scale we treat all quarks with masses less than that scale as 'active' but massless and we ignore quarks with masses greater than that scale. As a consequence, the coefficients k_n and u_n are now N_f-dependent and hence they will depend on the c.m. energy scale.

We also perform matching of the coupling at $Q^2 = m_b^2$. At LO and NLO this amounts to demanding continuity of the coupling at the threshold, but at NNLO and beyond, this

continuity is violated. The matching is governed by the decoupling theorem [188] and leads to the following relation between the coupling above and below a quark threshold (denoted by $a_{f+1}(Q^2)$ and $a_f(Q^2)$ respectively) [187, 189],

$$a_f(m_b^2) = a_{f+1}(m_b^2) + \frac{11}{72}\left(a_{f+1}(m_b^2)\right)^3. \tag{5.23}$$

This matching is implemented by adopting different values of the scale parameter in different N_f regions. This is governed by the following equations [187],

$$\Lambda_{N_f+1}^2 = \Lambda_{N_f}^2 \left(\frac{m_{N_f+1}^2}{\Lambda_{N_f}^2}\right)^{1-\frac{b^{N_f}}{b^{N_f+1}}} \times \exp\left(\frac{\delta_{NLO}+\delta_{NNLO}}{2b^{N_f+1}}\right), \tag{5.24}$$

where δ_{NLO} and δ_{NNLO} are given by

$$\delta_{NLO} = 4(c^{N_f+1}-c^{N_f})\ln L_m - 4c^{N_f+1}\ln\frac{b^{N_f+1}}{b^{N_f}}, \tag{5.25}$$

$$\begin{aligned}\delta_{NNLO} = {} & \frac{8}{b^{N_f}L_m}\Big(\left(c^{N_f+1}-c^{n_f}\right)c^{N_f}\ln L_m + \left(c^{N_f+1}\right)^2 - \left(c^{N_f}\right)^2 \\ & + c_2^{N_f} - c_2^{N_f+1} + \frac{7}{384}\Big).\end{aligned} \tag{5.26}$$

Here, Λ_{N_f} is the scale parameter in the region where N_f quarks are active, m_{N_f} is the pole mass of the f quark, b^{N_f}, c^{N_f} and $c_2^{N_f}$ are simply b, c and c_2 evaluated for N_f quark flavours and $L_m = \ln\left(m_{N_f+1}^2/\Lambda_{N_f}^2\right)$. The above relations ensure that Eq. (5.23) is satisfied by the three-loop coupling of Eq. (1.118). It must be noted that even for NLO predictions (where continuity of the coupling is assured) the observable is still discontinuous due to the coefficients k_n and u_n being N_f-dependent.

We must be careful how we apply this matching to the different results we have obtained. Equation (5.24) is an $\overline{\rm MS}$ result and hence we carry out the matching for $\Lambda_{\overline{\rm MS}}$ and then convert to the various other scales we have defined in Eqs. (5.4), (5.11) and (5.13). Also, $\tilde{\mathcal{K}}_{NLO}$ is an NLO result and hence, when applying Eq. (5.24) to this result, we omit δ_{NNLO}. The results in Eqs. (5.15), (5.19) and (5.22) are all at least NNLO and hence we use the full result from Eq. (5.24).

We wish to compare these results with relevant experimental data. Data is available for the GLS sum rule [97] and the polarized Bjorken sum rule [102, 181–183], where the

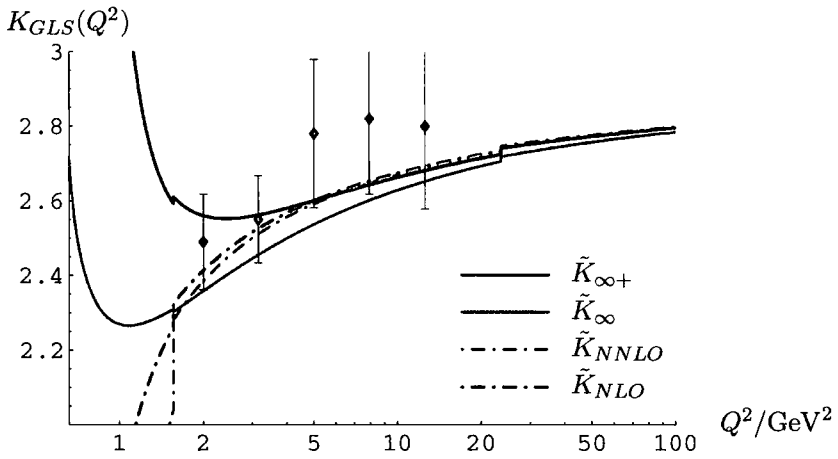


Figure 5.1: Various predictions for the GLS sum rule, superimposed on to experimental data.

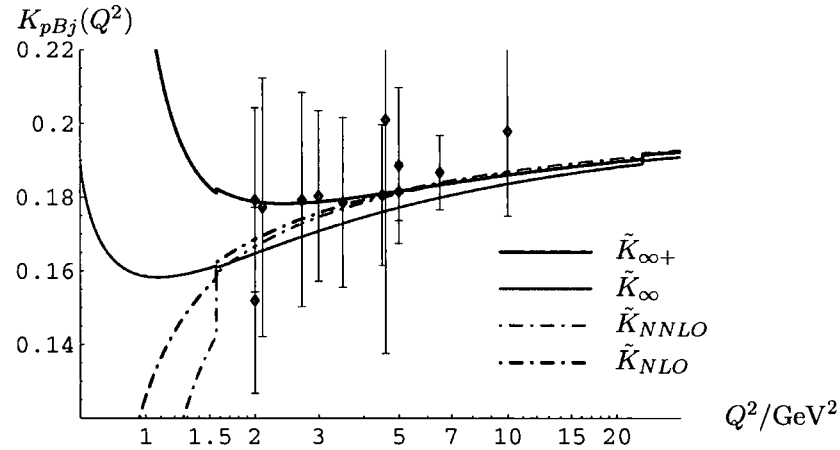


Figure 5.2: Various predictions for the polarized Bjorken sum rule, superimposed on to experimental data.

points we plot are those arising from the analysis of Ref. [184]. We plot the data in Figs. 5.1 and 5.2 along with our predictions for the full observables. No experiment has so far measured the unpolarized Bjorken sum rule, however the possibility exists that it may be extracted from experiments at a future neutrino factory [190]. We have taken $\Lambda^{(5)}_{\overline{\rm MS}} = 207$ MeV, corresponding to the world average value $\alpha_s(M_Z) = 0.1176$ [32].

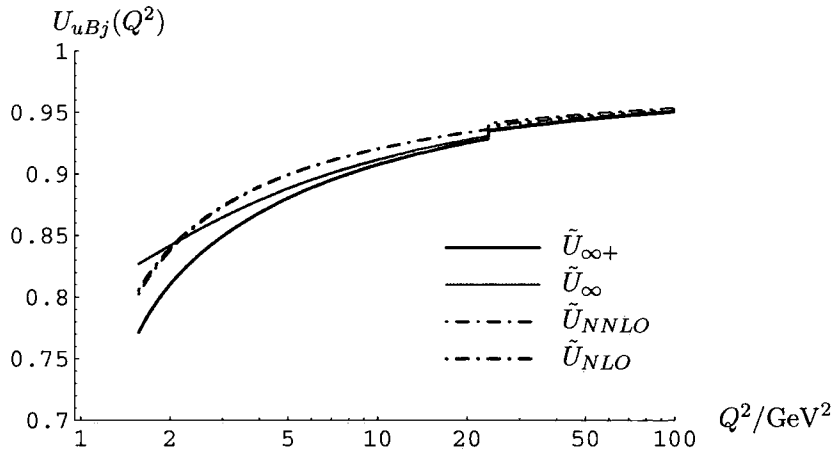


Figure 5.3: Various predictions for the unpolarized Bjorken sum rule.

Within the considerable error bars we see that the different versions of CORGI all show good agreement with the data. The important conclusion is that below $Q^2 \sim 5$ GeV^2 the all-orders and fixed-order approaches give drastically different predictions for the GLS and pBj sum rules. This indicates that fixed-order perturbation theory cannot be trusted for these lower energies. Even though we have no experimental data points for $U_{uBj}(Q^2)$ we have plotted the different versions of CORGI in Fig. 5.3. We see that fixed-order perturbation theory cannot be trusted below $Q^2 \sim 2$ GeV^2.

We can contrast these CORGI resummations with a leading-b resummation in the $\overline{\mathrm{MS}}$ scheme. If we choose the scale $\mu = xQ$ we have, analogous to Eq. (5.15),

$$\mathcal{K}_{\infty}^{\overline{\mathrm{MS}}}(x) \equiv a(xQ) + \sum_{n=1}^{\infty} k_n^{(n)}(x) b^n a^{n+1}(xQ) \tag{5.27}$$

$$= \mathcal{K}_{\infty}(a_v(x)). \tag{5.28}$$

Here $a(xQ)$ denotes the full higher-loop (three-loop if matching to NNLO) $\overline{\mathrm{MS}}$ coupling at scale $\mu = xQ$, and the $k_n(x)$ coefficients are the $\overline{\mathrm{MS}}$ coefficients with scale $\mu = xQ$. The coupling $a_v(x)$ can be related to $a(xQ)$ using Eqs. (1.133) and (5.14),

$$\frac{1}{a_v(x)} = \frac{1}{a(xQ)} - b\left(\ln x + \frac{5}{6}\right). \tag{5.29}$$

We can match this all-orders resummation to the exact NLO and NNLO perturbative coefficients obtaining, in analogy with Eq. (5.19),

$$\mathcal{K}^{\overline{\rm MS}}_{\infty+}(x) \;=\; \mathcal{K}^{\overline{\rm MS}}_{\infty}(x) + (k_1(x) - bk_1^{(1)}(x))a^2(xQ) + (k_2(x) - b^2k_2^{(2)}(x))a^3(xQ). \quad (5.30)$$

In this equation, $a(Q)$ is the approximated three-loop coupling, given in Eq. (1.118) [32].

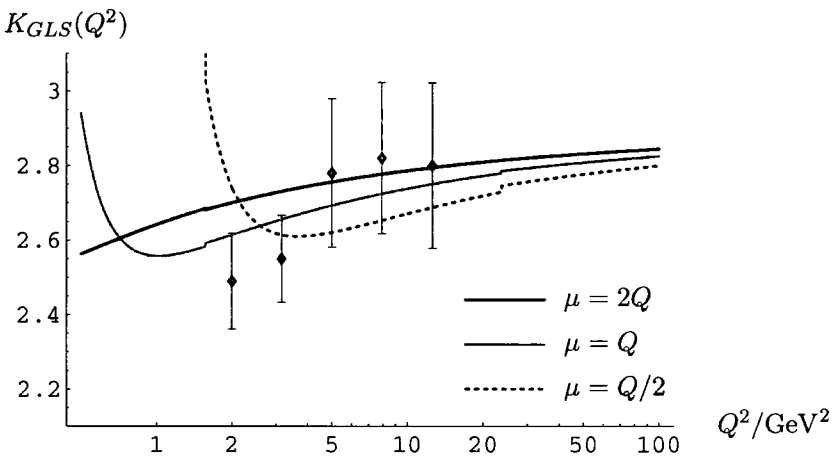


Figure 5.4: All-orders leading-b prediction for the GLS sum rule. Here we use the $\overline{\rm MS}$ result of Eq. (5.30) with renormalization scales of $\mu = 2Q, Q$ and $Q/2$.

The resulting prediction is plotted in Fig. 5.4 for the GLS sum rule. Three different matching scales corresponding to $x = 2, 1, \frac{1}{2}$ were chosen. As can be seen, the matched resummed perturbative result is hopelessly x-dependent. The CORGI result $\tilde{\mathcal{K}}_{\infty+}(Q^2)$ corresponds to the choice $x = e^{-k_1(x=1)/b}$. Notice that no NLO matching is required in the CORGI approach, since an infinite set of RG-predictable terms involving k_1 have been resummed to all-orders, yielding a μ-independent result cf. Eq. (1.166).

5.4 Non-perturbative corrections

In addition to perturbative corrections it is expected that there will be non-perturbative, Higher-Twist (HT) corrections to the sum rules. These were discussed in the previous chapter in the language of the one-chain term of the skeleton expansion. The leading-b resummations we have discussed can be written in the form derived from

Eq. (4.6),

$$\mathcal{K}_\infty(Q^2) = \int_0^1 dt\, \omega_{\mathcal{K}}^{\rm IR}(t) a(tQ^2) + \int_1^\infty dt\, \omega_{\mathcal{K}}^{\rm UV}(t) a(tQ^2). \tag{5.31}$$

Here $\omega_{\mathcal{K}}(t)$ is the characteristic function which is piecewise continuous at $t = 1$. As detailed in chapter 4, the two terms respectively reproduce the IR renormalon and UV renormalon contributions in the Borel sum. The skeleton expansion result is also well-defined for $Q^2 < \Lambda^2$ where the standard Borel representation breaks down.

By considering the compensation of ambiguities between perturbative and non-perturbative corrections we are led to an expression for HT corrections in terms of the characteristic function,

$$\mathcal{K}_{\rm HT}(Q^2) = \kappa \frac{\Lambda^2}{Q^2} \omega_{\mathcal{K}}^{\rm IR}\left(\frac{\Lambda^2}{Q^2}\right), \tag{5.32}$$

(see Eqs. (4.82) - (4.84)). Here κ is a single overall non-perturbative constant. This result is obtained from the V-scheme result of Eq. (4.32) and hence the appropriate scale parameter for Eq. (5.32) is $\Lambda_{\rm V}$. A similar expression holds for $\mathcal{U}_{\rm HT}(Q^2)$ in terms of $\omega_{\mathcal{U}}^{\rm IR}(t)$.

The characteristic functions $\omega_{\mathcal{K}}(t)$ and $\omega_{\mathcal{U}}(t)$ can be determined from their Borel representations (Eq. (4.10) in conjunction with Eqs. (4.27) and (4.28)) by reversing the changes of variables used to obtain Eq. (4.59). $\omega_{\mathcal{K}}(t)$ has the form,

$$\omega_{\mathcal{K}}^{\rm IR}(t) = \frac{8}{9} - \frac{5}{9}t, \tag{5.33}$$

$$\omega_{\mathcal{K}}^{\rm UV}(t) = \frac{4}{9t^2} - \frac{1}{9t^3}, \tag{5.34}$$

and the equivalent expression for $\mathcal{U}(a)$ is,

$$\omega_{\mathcal{U}}^{\rm IR}(t) = \frac{4}{3} - t, \tag{5.35}$$

$$\omega_{\mathcal{U}}^{\rm UV}(t) = \frac{1}{3t^3}. \tag{5.36}$$

Each term in these expansions is in one-to-one correspondence with a renormalon singularity in the Borel transforms of Eqs. (4.27) and (4.28).

In Figs. 5.5 and 5.6 we take the four expressions in Eqs. (5.15), (5.19), (5.21) and

	κ_{GLS}	$-3\left(\frac{8}{9}\right)\kappa_{GLS}\,\Lambda_V^2/\mathrm{GeV}^2$	χ^2_{GLS}/d.o.f.
$\tilde{\mathcal{K}}_{NLO}$	-0.166 ± 0.20	0.1007 ± 0.12	1.256/4
$\tilde{\mathcal{K}}_{NNLO}$	-0.217 ± 0.20	0.1313 ± 0.12	1.238/4
$\tilde{\mathcal{K}}_{\infty}$	-0.33 ± 0.20	0.1997 ± 0.12	1.804/4
$\tilde{\mathcal{K}}_{\infty+}$	0.0216 ± 0.20	-0.01311 ± 0.12	2.089/4

Table 5.1: Fitted values of the non-perturbative constant κ_{GLS}, together with their respective χ^2 per degree of freedom (d.o.f).

	κ_{pBj}	$-\frac{1}{6}\left\|\frac{g_A}{g_V}\right\|\left(\frac{8}{9}\right)\kappa_{pBj}\,\Lambda_V^2/\mathrm{GeV}^2$	χ^2_{pBj}/d.o.f.
$\tilde{\mathcal{K}}_{NLO}$	-0.06255 ± 0.26	0.002647 ± 0.011	1.173/11
$\tilde{\mathcal{K}}_{NNLO}$	-0.1085 ± 0.26	0.004593 ± 0.011	1.162/11
$\tilde{\mathcal{K}}_{\infty}$	-0.2823 ± 0.26	0.01195 ± 0.011	1.456/11
$\tilde{\mathcal{K}}_{\infty+}$	0.03957 ± 0.26	-0.001674 ± 0.011	1.662/11

Table 5.2: Fitted values of the non-perturbative constant κ_{pBj}, together with their respective χ^2 per degree of freedom.

(5.22), supplemented by the non-perturbative term in Eq. (5.32), and perform fitting to experimental data for both the GLS and the polarized Bjorken sum rules separately. We fix $\Lambda^{(5)}_{\overline{\rm MS}} = 207$ MeV, as before, and use χ^2 fitting to obtain the optimal value of the non-perturbative parameters[3] in each case, κ_{GLS} and κ_{pBj}. The fitted parameters are summarized in tables 5.1 and 5.2. We note that the coefficient of a $1/Q^2$ (twist-4) power correction to either Eq. (4.23) or Eq. (4.22) corresponds to a value of $-8\Lambda_V^2\kappa_{GLS}/3$ or $-\frac{4}{27}|g_A/g_V|\Lambda_V^2\kappa_{pBj}$, respectively. The corresponding values in GeV^2 are presented in tables 5.1 and 5.2. For comparison, the central values resulting from a three-point function QCD sum rules fit [91–93] are -0.294 and -0.013, respectively. We see that the power corrections in tables 5.1 and 5.2 resulting from fitting to the $\tilde{\mathcal{K}}_{\infty+}$ all-orders resummations are significantly smaller, although with large error bars. Connections between power corrections for the three DIS sum rules have also been explored in Ref. [191].

[3]We assume that the non-perturbative parameters in the cases of K_{GLS} and K_{pBj} are independent of each other.

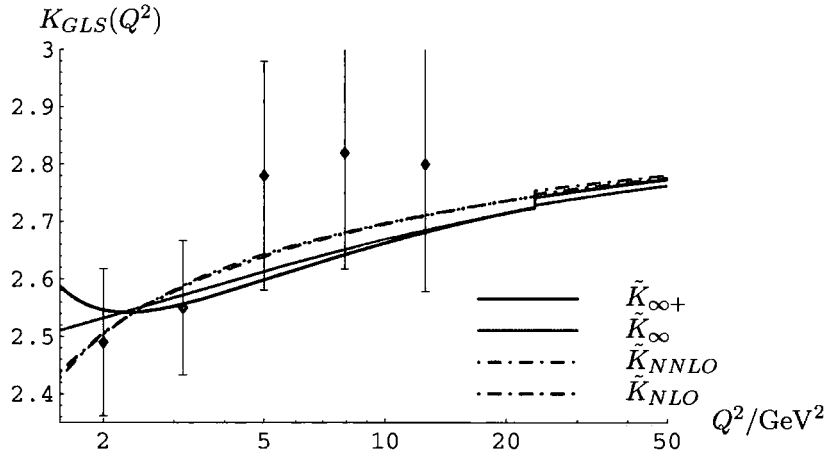


Figure 5.5: Predictions for the GLS sum rule, including both perturbative and non-perturbative corrections, fitted to the data by varying the parameter κ_{GLS}.

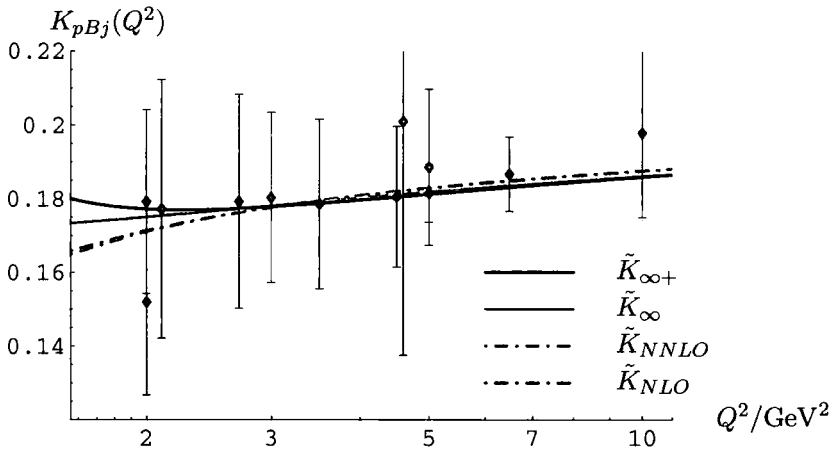


Figure 5.6: Predictions for the polarized Bjorken sum rule, including both perturbative and non-perturbative corrections, fitted to the data by varying the parameter κ_{pBj}.

5.5 Summary and conclusions

By comparing with leading-b resummations based on the CORGI approach we attempted to infer the validity of fixed-order perturbation theory for the GLS and Bjorken DIS sum rules at small energy scales $Q^2 \sim 3$ GeV^2, where data has been used to extract $\alpha_s(M_Z)$ [32]. Figures 5.1 and 5.2 indicate that fixed-order perturbation theory and re-

summed all-orders predictions start to differ drastically for energies below $Q^2 \sim 5$ GeV^2 in the GLS and pBj cases, and for energies below $Q^2 \sim 2$ GeV^2 in the uBj case.

The use of the CORGI approach ensures that all RG-predictable scale-dependent logarithms are resummed to all-orders, avoiding the need for NLO matching, which would otherwise make the conclusions about the validity of fixed-order perturbation theory dependent on the chosen matching scale, as illustrated in Fig. 5.4.

We performed fits to data for power corrections using the model proposed in chapter 4 and based on the skeleton expansion characteristic function. The size of power corrections inferred by fitting the all-orders CORGI resummations was much smaller than the central values obtained from three-point QCD sum rule estimates of Refs. [91–93].

Chapter 6

Analysis of moments of F_3 in neutrino-nucleon scattering

In this chapter we turn our attention to fixed-order predictions for moments of DIS structure functions. We present an analysis of the moments of F_3 in neutrino-nucleon scattering, our aim being to test the accuracy of QCD predictions by comparing them with experimental data. In particular, we are interested in how this analysis is affected by the removal of renormalization and factorization scale dependence via the CORGI approach to perturbation theory.

6.1 Introduction

Because of asymptotic freedom, we can apply perturbation theory to the dynamics of quarks at high energies. When supplemented by the OPE and the factorization procedure, we have at our disposal a powerful method of investigating the behaviour of hadron-bound quarks, via lepton-hadron scattering. In particular, factorization allows us to avoid the explicit calculation of non-perturbative effects, and hence allows us to obtain a result written in terms of perturbative variables, supplemented by a set of non-perturbative constants.

Despite this, there remain several obstacles which limit both the validity and accuracy of factorized predictions when applied to experimentally measured structure functions. Two of these are outlined below:

- QCD predictions for the moments are both factorization and renormalization scheme (FRS) dependent. This is effectively a double dose of the problem of RS dependence discussed in section 1.8. The canonical approach to dealing with this is to set the arbitrary scales equal to some 'physical scale' i.e. $M = \mu = Q$, and then to adopt some 'best scheme' (such as $\overline{\text{MS}}$) for making the finite subtractions in each of the cases of renormalization and factorization. However, the resultant expressions then suffer from dependence on these arbitrary choices, and therefore have an inherent theoretical error associated with them. Moreover, there is a lack of physical motivation to justify any particular scheme being the 'best' choice.

- There are certain limitations on the ranges of Q^2 and x (defined in Eqs. (2.2) - (2.4)) for which data for the structure functions is available. In particular, there are upper and lower limits on the accessible range of x at low and high Q^2 respectively. Experimental values for the moments at a given Q^2 are obtained by integration of the structure function (measured *at* that Q^2), over the full range of x. In order to do this, it is necessary to model the available structure function data over this x-range.

  However, regions of x for which we have no data will leave the structure function modelling (and therefore the experimental moments) poorly constrained. As a consequence, the result of any analysis will be dependent on the way the structure functions are modelled (in the missing data regions). This is clearly undesirable and we are forced to either live with this modelling dependency, or expand the errors on the experimental values for the moments accordingly.

In this chapter, we intend to overcome these problems and to improve upon previous analyses in the following ways:

- We use the method of Bernstein averages to overcome the above-mentioned limitations of the structure function data. This technique was first applied in Refs. [192,193], and was adapted for CORGI analysis in Ref. [194]. In this method, data for a structure function at a given Q^2 is multiplied by a polynomial in x (which is peaked sharply within the region for which we have data) and then integrated over the full range of x. The result is a *Bernstein average* – an object

which has negligible dependence on the values of the structure functions in the missing data regions, and can be written as a linear combination of moments. Thus, by comparing theoretical and experimental values of the Bernstein averages (rather than the moments themselves) we can reduce any dependence of the analysis on the way we model the structure functions in the inaccessible regions of x.

- Until recently, only partial results for the NNLO coefficients of the anomalous dimension were available. This has limited any analysis of F_3 to the subset of odd moments $n = 1, 3, 5, \ldots 13$. However, due to the results in Ref. [78], we now have analytic expressions for the non-singlet anomalous dimension for any value of n. Hence, we no longer need to limit ourselves to odd moments, and we can also remove the limit on the highest moment available to the analysis.

- We will adapt the CORGI approach (previously applied to the elimination of renormalization scale dependence, in section 1.8.3) to the problem of FRS dependence (in which we have *two* arbitrary scales), and thus reduce any dependence of the analysis on renormalization or factorization scales. This approach was previously outlined in Refs. [39,194,195]. However, the expressions for the explicit FRS dependence of the moments derived in those works were incorrect. We correct these results and apply them to the analysis of moments of F_3.

The material in this chapter is derived from the research carried out in Ref. [196], and is organized as follows:

First, we describe how the FRS dependence of moments can be parameterized and how, as a result of this, the CORGI approach to perturbation theory may be generalized to deal with dependence on *two* arbitrary scales (M and μ). We show how the explicit FRS dependence of the moments may be obtained, before deriving perturbative predictions for them in the CORGI, physical scale and effective charge approaches. We then describe how these predictions are extended to include quark mass threshold, target mass, and higher twist effects. Next, we describe how the method of Bernstein averages allows us to overcome limitations in the data for F_3 and how this approach can be adapted to deal with both odd *and* even moments. Finally we present the results of our analysis and discuss what conclusions we can draw from them.

6.2 Factorization and renormalization scheme dependence of F_3 moments

The moments we are concerned with in this chapter are those derived from F_3 in (anti)neutrino-nucleon scattering. They are defined as:

$$\mathcal{M}_3^{\nu N}(n;Q^2) \;=\; \int_0^1 dx\; x^{n-1} F_3^{\nu N}(x,Q^2). \tag{6.1}$$

As stated in section 2.7.2, these moments can be factorized in the following form,

$$\mathcal{M}_3^{\nu N}(n;Q^2) \;=\; \langle N|\mathcal{O}_{n,\,\mathrm{NS}}(M)|N\rangle\, C_n^{(3)}(Q,M,\mu,a(\mu)), \tag{6.2}$$

where $\langle N|\mathcal{O}_{n,\,\mathrm{NS}}(M)|N\rangle$ is the NS operator matrix element of nucleon states and $C_n^{(3)}(Q,M,\mu,a(\mu))$ is the coefficient function. The operator matrix element is factorized at the scale M into a non-perturbative component and a perturbative expression. The perturbative component is written in terms of the coupling evaluated at the factorization scale, $a \equiv a(M)$. The factorization scale dependence is then governed by the anomalous dimension equation,

$$M\frac{\partial}{\partial M}\langle N|\mathcal{O}_{n,\,\mathrm{NS}}(M)|N\rangle \;=\; \langle N|\mathcal{O}_{n,\,\mathrm{NS}}(M)|N\rangle\, \gamma_{n,\,\mathrm{NS}}(a), \tag{6.3}$$

where $\gamma_{n,\,\mathrm{NS}}(a)$ is the non-singlet anomalous dimension and has the following perturbative expansion,

$$\gamma_{n,\,\mathrm{NS}}(a) \;=\; -d(n)a - d_1(n)a^2 - d_2(n)a^3 - \ldots. \tag{6.4}$$

The M dependence of the coupling is governed by the beta-function equation,

$$M\frac{\partial a}{\partial M} \;=\; -ba^2(1 + ca + c_2a^2 + \ldots). \tag{6.5}$$

These two equations determine the perturbative behaviour of the operator matrix element. For the remainder of this chapter, we simplify our notation by dropping the sub- and superscripts 'νN', 'n', '(3)' and 'NS', from the quantities in Eqs. (6.2) and (6.3). Also, although the coefficients $d_i(n)$ in Eq. (6.4) are n-dependent we suppress this, again for the sake of brevity. In section 1.8.1 we described how we can parameterize RS dependence in the case where we have a single arbitrary scale. In the case of the moments however, we have two arbitrary scales and hence we must adapt the approach in order to take this into account.

A solution to Eqs. (6.3) and (6.5) can be obtained in the form,

$$\langle \mathcal{O}(M) \rangle = A_n \exp\left\{ \int_0^{a(M)} \frac{\gamma(x)}{\beta(x)} dx - \int_0^{\infty} \frac{\gamma^{(1)}(x)}{\beta^{(2)}(x)} dx \right\}, \tag{6.6}$$

where $\gamma^{(i)}$ and $\beta^{(i)}$ denote the anomalous dimension and beta-function equation truncated after i terms, and we have defined $\langle \mathcal{O}(M) \rangle \equiv \langle N | \mathcal{O}_{n,\,\mathrm{NS}}(M) | N \rangle$. There is a distinct parallel between the above equation and the solution to the beta function in Eqs. (1.110) and (1.111). The second integral in Eq. (6.6) is an infinite constant. We are free to choose any form we wish for this term, subject to the constraint that it must have the same singularity structure as the first term in the integrand. However, a particular choice for this constant corresponds to a particular definition of A_n. Consequently, A_n can be likened to the dimensional transmutation parameter, Λ, in that it defines the missing boundary condition in Eq. (6.3). A_n is actually a (set of) non-perturbative constant(s), generated by the factorization process. Their precise values cannot be calculated within perturbation theory, and hence must be obtained by comparison with experimental data.

The coefficient function $C_n(Q, M, \mu, a(\mu))$ depends on both the renormalization *and* factorization scheme adopted, and it takes the form of an expansion in powers of the coupling evaluated at the renormalization scale,

$$C_n(Q, M, \mu, a(\mu)) = 1 + r_1 \tilde{a} + r_2 \tilde{a}^2 + \ldots, \tag{6.7}$$

where $\tilde{a} \equiv a(M = \mu)$. Using the above equation together with Eqs. (6.2) and (6.6), the moments can be written as [39],

$$\mathcal{M}(n; Q^2) = A_n \left(\frac{ca}{1+ca} \right)^{d/b} \exp(\mathcal{I}(a))(1 + r_1 \tilde{a} + r_2 \tilde{a}^2 + \cdots), \tag{6.8}$$

where,

$$\mathcal{I}(a) = \int_0^a dx \frac{d_1 + (d_1 c + d_2 - d c_2)x + (d_3 + c d_2 - c_3 d)x^2 + \cdots)}{b(1+cx)(1 + cx + c_2 x^2 + c_3 x^3 + \cdots)}. \tag{6.9}$$

The explicit M dependence of the coupling can be obtained by solving the following transcendental equation [33],

$$\frac{1}{a} + c \ln \frac{ca}{1+ca} = b \ln \frac{M}{\tilde{\Lambda}} - b \int_0^a \left[\frac{1}{\beta(x)} - \frac{1}{\beta^{(2)}(x)} \right]. \tag{6.10}$$

Equation (6.8) serves as a prototypical expression for the moments, from which we derive perturbative expansions for them using three different approaches described in sections 1.8.2 - 1.8.4. However, we must first adapt these approaches to the two-scale case.

6.2.1 Parameterizing FRS dependence

In the same way that the beta-function equation governs the RS dependence of the coupling, the anomalous dimension equation governs the FS dependence of the operator matrix element. Hence the FRS dependence of the moments is characterized by the coefficients of these two equations. The first coefficient of the anomalous dimension is FS invariant and therefore it is the set of parameters $\{M, \mu, c_2, c_3, \ldots, d_1, d_2, \ldots\}$ that are relevant to FRS dependence [103]. Moreover, these parameters can be used to define any particular FRS and also parameterize the FRS dependence of the moments.

The self-consistency of perturbation theory demands that the FRS dependence of an $\mathcal{O}(a^N)$ result, $\mathcal{M}^{(N)}(n; Q^2)$, is at least of order $N+1$, hence formally,

$$\frac{\partial \mathcal{M}^{(N)}}{\partial(\mathcal{FRS})} = \mathcal{O}(a^{N+1}). \tag{6.11}$$

Here, as in Eq. (1.120), we use $\mathcal{FRS}$ to symbolize the set of parameters defining a particular scheme.

Inspired by the parameterization of RS dependence in section 1.8.1, we now proceed to systematically determine the explicit FRS dependence of the coefficients r_i. This is achieved by differentiating the expressions for the moments with respect to the FRS parameters, $\{M, \mu, c_2, \ldots, d_1, d_2, \ldots\}$ [103], and in order to do this, we will need the following results.

We can write $\mathcal{I}(a)$ as an expansion in powers of a,

$$\mathcal{I}(a) = \mathcal{I}_1 a + \mathcal{I}_2 a^2 + \mathcal{I}_3 a^3 + \mathcal{O}(a^4), \tag{6.12}$$

where the coefficients are,

$$\mathcal{I}_1 = \frac{d_1}{b}, \tag{6.13}$$

$$\mathcal{I}_2 = \frac{d_2}{2b} - \frac{cd_1}{2b} - \frac{c_2 d}{2b}, \tag{6.14}$$

$$\mathcal{I}_3 = \frac{d_3}{3b} - \frac{d_2 c}{3b} - \frac{c_3 d}{3b} - \frac{c_2 d_1}{3b} + \frac{2c_2 dc}{3b} + \frac{d_1 c^2}{3b}. \tag{6.15}$$

Also, using the beta-function equation, we can write the coupling $\tilde{a}$ as an expansion in powers of a,

$$\tilde{a} = a + La^2 + (L^2 + cL)a^3 + \mathcal{O}(a^4), \tag{6.16}$$

and consequently,

$$\tilde{a}^2 = a^2 + 2La^3 + \mathcal{O}(a^4), \tag{6.17}$$

where we have defined,

$$L = b \ln[M/\mu]. \tag{6.18}$$

Factorization scale dependence

Differentiating Eq. (6.8) with respect to $\ln M$ gives,

$$M\frac{\partial \mathcal{M}}{\partial M} = \mathcal{M}\left[\frac{1+ca}{ca}\left(\frac{c}{1+ca} - \frac{c^2 a}{(1+ca)^2}\right)\frac{d}{b}\beta(a) + (\mathcal{I}_1 + 2\mathcal{I}_2 a + 3\mathcal{I}_3 a^2)\beta(a) + \frac{\tilde{a}M\frac{\partial r_1}{\partial M} + \tilde{a}^2 M\frac{\partial r_2}{\partial M} + \tilde{a}^3 M\frac{\partial r_3}{\partial M}}{1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3} + \ldots\right], \tag{6.19}$$

where $\mathcal{M} \equiv \mathcal{M}(n; Q^2)$. Using Eqs. (6.12) - (6.18), and then expanding in powers of a gives,

$$\begin{aligned} M\frac{\partial \mathcal{M}}{\partial M} = \mathcal{M}\Bigg[& -da - d_1 a^2 - d_2 a^3 + aM\frac{\partial r_1}{\partial M} + a^2\left(M\frac{\partial r_2}{\partial M} + M\frac{\partial r_1}{\partial M}(L - r_1)\right) \\ & + a^3\left(M\frac{\partial r_3}{\partial M} + M\frac{\partial r_2}{\partial M}\left(2L - r_1\right)\right. \\ & \left. + M\frac{\partial r_1}{\partial M}\left(r_1^2 - r_2 - 2r_1 L + L^2 + cL\right)\right) + \mathcal{O}\left(a^4\right)\Bigg]. \end{aligned} \tag{6.20}$$

From Eq. (6.11) we know that the term in square brackets in the above equation must vanish order-by-order in a. Thus, equating the coefficients of powers of a, and then setting to zero yields,

$$M\frac{\partial r_1}{\partial M} = d, \tag{6.21}$$

$$M\frac{\partial r_2}{\partial M} = dr_1 - dL + d_1, \tag{6.22}$$

$$M\frac{\partial r_3}{\partial M} = d_2 + d_1 r_1 + dr_2 - dr_1 L - 2d_1 L + dL^2 - dcL. \tag{6.23}$$

Renormalization scale dependence

The $\log\mu$ dependence is somewhat easier to determine. Differentiating $\mathcal{M}(n;Q^2)$ with respect to $\log\mu$ gives,

$$\mu\frac{\partial \mathcal{M}}{\partial \mu} = \frac{\mathcal{M}}{1 + r_1\tilde{a} + r_2\tilde{a}^2}\left[\mu\frac{\partial r_1}{\partial \mu}\tilde{a} + r_1\beta(\tilde{a}) + \mu\frac{\partial r_2}{\partial \mu}\tilde{a}^2 + 2r_2\tilde{a}\beta(\tilde{a}) + \mu\frac{\partial r_3}{\partial \mu}\tilde{a}^3 + 3r_3\tilde{a}^2\beta(\tilde{a}) + \mathcal{O}\left(\tilde{a}^4\right)\right]. \tag{6.24}$$

Comparing orders of $\tilde{a}$ is equivalent to comparing orders of a and therefore, we can obtain,

$$\mu\frac{\partial r_1}{\partial \mu} = 0, \tag{6.25}$$

$$\mu\frac{\partial r_2}{\partial \mu} = r_1 b, \tag{6.26}$$

$$\mu\frac{\partial r_3}{\partial \mu} = 2r_2 b + b r_1 c. \tag{6.27}$$

Dependence on d_i

First, for d_1 we have,

$$\frac{\partial \mathcal{M}}{\partial d_1} = \mathcal{M}\left[\frac{\partial \mathcal{I}(a)}{\partial d_1} + \frac{\tilde{a}\frac{\partial r_1}{\partial d_1} + \tilde{a}^2\frac{\partial r_2}{\partial d_1} + \tilde{a}^3\frac{\partial r_3}{\partial d_1}}{1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3} + \ldots\right], \tag{6.28}$$

and expanding this in powers of a gives,

$$\frac{\partial \mathcal{M}}{\partial d_1} = \mathcal{M}\left[\frac{1}{b}\left(a - \frac{ca^2}{2} + \frac{a^3}{3}\left(c^2 - c_2\right)\right) + a\frac{\partial r_1}{\partial d_1} + a^2\left(\frac{\partial r_2}{\partial d_1} + \frac{\partial r_1}{\partial d_1}(L - r_1)\right)\right.$$
$$\left. + a^3\left(\frac{\partial r_3}{\partial d_1} + \frac{\partial r_2}{\partial d_1}\left(2L - r_1\right) + \frac{\partial r_1}{\partial d_1}\left(r_1^2 - r_2 - 2r_1L + L^2 + cL\right)\right) + \mathcal{O}\left(a^4\right)\right]. \tag{6.29}$$

Equating coefficients then yields,

$$\frac{\partial r_1}{\partial d_1} = -\frac{1}{b}, \tag{6.30}$$

$$\frac{\partial r_2}{\partial d_1} = \frac{c}{2b} - \frac{r_1}{b} + \frac{L}{b}, \tag{6.31}$$

$$\frac{\partial r_3}{\partial d_1} = \frac{cr_1}{2b} - \frac{c^2}{3b} - \frac{r_2}{b} + \frac{c_2}{3b} - \frac{L^2}{b} + \frac{Lr_1}{b}. \tag{6.32}$$

Performing the same procedure for d_2 gives,

$$\frac{\partial \mathcal{M}}{\partial d_2} = \mathcal{M}\left[\frac{\partial \mathcal{I}(a)}{\partial d_2} + \frac{\tilde{a}\frac{\partial r_1}{\partial d_2} + \tilde{a}^2\frac{\partial r_2}{\partial d_2} + \tilde{a}^3\frac{\partial r_3}{\partial d_2}}{1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3} + \ldots\right]$$
$$= \mathcal{M}\left[\frac{a^2}{2b} - \frac{ca^3}{3b} + a\frac{\partial r_1}{\partial d_2} + a^2\left(\frac{\partial r_2}{\partial d_2} + \frac{\partial r_1}{\partial d_2}(L - r_1)\right)\right.$$
$$\left. + a^3\left(\frac{\partial r_3}{\partial d_2} + \frac{\partial r_2}{\partial d_2}\left(2L - r_1\right) + \frac{\partial r_1}{\partial d_2}\left(r_1^2 - r_2 - 2r_1L + L^2 + cL\right)\right) + \mathcal{O}\left(a^4\right)\right], \tag{6.33}$$

and equating coefficients yields,

$$\frac{\partial r_1}{\partial d_2} = 0, \tag{6.34}$$

$$\frac{\partial r_2}{\partial d_2} = -\frac{1}{2b}, \tag{6.35}$$

$$\frac{\partial r_3}{\partial d_2} = \frac{c}{3b} + \frac{1}{2b}(2L - r_1). \tag{6.36}$$

Finally, the only term relevant to the d_3 dependence is the following differential,

$$\frac{\partial \mathcal{I}(a)}{\partial d_3} = \frac{a^3}{3b}, \tag{6.37}$$

and therefore we infer that,

$$\frac{\partial r_3}{\partial d_3} = -\frac{1}{3b}. \tag{6.38}$$

Dependence on c_i

Finally, we turn our attention to the dependence of the moments on the RS dependent coefficients of the beta-function equation, c_i. For this we will need the results of Eq. (1.138),

$$\frac{\partial a}{\partial c_2} = a^3 + \mathcal{O}\left(a^5\right), \qquad \frac{\partial a}{\partial c_3} = \frac{a^4}{2} + \mathcal{O}\left(a^5\right). \tag{6.39}$$

Also, to simplify things, we state the differentials of $\tilde{a}$ and $\tilde{a}^2$ with respect to c_2,

$$\frac{\partial \tilde{a}}{\partial c_2} = \frac{\partial \tilde{a}}{\partial a}\frac{\partial a}{\partial c_2} \tag{6.40}$$

$$= (1 + 2La + 3(L^2 + cL)a^2 + \cdots)(a^3 + \cdots) \tag{6.41}$$

$$= a^3 + 2La^4 + \mathcal{O}\left(a^5\right), \tag{6.42}$$

and for $\tilde{a}^2$, we have,

$$\frac{\partial \tilde{a}^2}{\partial c_2} = \frac{\partial \tilde{a}^2}{\partial a}\frac{\partial a}{\partial c_2} \tag{6.43}$$

$$= (2a + 6La^2 + \cdots)(a^3 + \cdots) \tag{6.44}$$

$$= 2a^4 + \mathcal{O}\left(a^5\right). \tag{6.45}$$

The derivative of $\tilde{a}^3$ with respect to c_2 is of minimum order a^5 and hence can be neglected.

Differentiating $\mathcal{M}(n; Q^2)$ with respect to c_2 gives,

$$\frac{\partial \mathcal{M}}{\partial c_2} = \mathcal{M}\left[\frac{d}{b}\left(\frac{1+ca}{ca}\right)\left(\frac{c}{1+ca} - \frac{c^2 a}{(1+ca)^2}\right)\frac{\partial a}{\partial c_2} + \frac{\partial \mathcal{I}(a)}{\partial c_2}\right.$$

$$+ \quad \frac{\tilde{a}\frac{\partial r_1}{\partial c_2} + \tilde{a}^2\frac{\partial r_2}{\partial c_2} + \tilde{a}^3\frac{\partial r_3}{\partial c_2}}{1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3} + \frac{r_1\frac{\partial \tilde{a}}{\partial c_2} + r_2\frac{\partial \tilde{a}^2}{\partial c_2} + r_3\frac{\partial \tilde{a}^3}{\partial c_2}}{1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3} + \ldots \Bigg]$$

$$= \mathcal{M}\Bigg[\frac{d}{b}\left(\frac{1}{a} - \frac{c}{1+ca}\right)a^3 - \frac{da^2}{2b} + a^3\frac{2cd - d_1}{3b} + \frac{\partial \mathcal{I}(a)}{\partial a}\frac{\partial a}{\partial c_2}$$

$$+ \quad \frac{\tilde{a}\frac{\partial r_1}{\partial c_2} + \tilde{a}^2\frac{\partial r_2}{\partial c_2} + \tilde{a}^3\frac{\partial r_3}{\partial c_2}}{1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3} + r_1a^3 + \mathcal{O}\left(a^4\right)\Bigg]$$

$$= \mathcal{M}\Bigg[\frac{da^2}{2b} + a^3\left(\frac{2d_1}{3b} - \frac{cd}{3b}\right) + a\frac{\partial r_1}{\partial c_2} + a^2\left(\frac{\partial r_2}{\partial c_2} + \frac{\partial r_1}{\partial c_2}(L - r_1)\right)$$

$$+ a^3\left(\frac{\partial r_3}{\partial c_2} + \frac{\partial r_2}{\partial c_2}\left(2L - r_1\right)\right.$$

$$\left. + \frac{\partial r_1}{\partial c_2}\left(r_1^2 - r_2 - 2r_1L + L^2 + cL\right) + r_1\right) + \mathcal{O}\left(a^4\right)\Bigg], \tag{6.46}$$

and equating coefficients yields,

$$\frac{\partial r_1}{\partial c_2} = 0, \tag{6.47}$$

$$\frac{\partial r_2}{\partial c_2} = -\frac{d}{2b}, \tag{6.48}$$

$$\frac{\partial r_3}{\partial c_2} = -\frac{r_1 d}{2b} + \frac{Ld}{b} + \frac{cd}{3b} - \frac{2d_1}{3b} - r_1. \tag{6.49}$$

In the case of c_3, we have,

$$\frac{\partial \mathcal{M}}{\partial c_3} = \mathcal{M}\Bigg[\frac{d}{b}\left(\frac{1}{a} - \frac{c}{1+ca}\right)\frac{a^4}{2} + \frac{\partial \mathcal{I}(a)}{\partial c_3} + \frac{\tilde{a}\frac{\partial r_1}{\partial c_3} + \tilde{a}^2\frac{\partial r_2}{\partial c_3} + \tilde{a}^3\frac{\partial r_3}{\partial c_3}}{1 + r_1\tilde{a} + r_2\tilde{a}^2 + r_3\tilde{a}^3} + \ldots\Bigg]$$

$$= \mathcal{M}\Bigg[\frac{da^3}{6b} + a\frac{\partial r_1}{\partial c_3} + a^2\left(\frac{\partial r_2}{\partial c_3} + \frac{\partial r_1}{\partial c_3}(L - r_1)\right)$$

$$+ \quad a^3\left(\frac{\partial r_3}{\partial c_3} + \frac{\partial r_2}{\partial c_3}\left(2L - r_1\right) + \frac{\partial r_1}{\partial c_3}\left(r_1^2 - r_2 - 2r_1L + L^2 + cL\right)\right) + \mathcal{O}\left(a^4\right)\Bigg], \tag{6.50}$$

and equating coefficients in this case, gives,

$$\frac{\partial r_1}{\partial c_3} = 0, \tag{6.51}$$

$$\frac{\partial r_2}{\partial c_3} = 0, \tag{6.52}$$

$$\frac{\partial r_3}{\partial c_3} = -\frac{d}{6b}. \tag{6.53}$$

FRS invariance of A_n

In the above derivations we have assumed that A_n is scheme invariant. This was shown in Ref. [197], by means of the following argument. Using the beta-function equation we can relate the coupling defined in two different schemes a and $\bar{a}$ (see Eq. (1.125)). We can also do this for the operator matrix element, using the anomalous dimension. Thus,

$$a(\mu) = \bar{a}(\mu)(1 + \nu_1\bar{a}(\mu) + \nu_2\bar{a}(\mu)^2 + \cdots), \tag{6.54}$$

$$\langle\mathcal{O}(M)\rangle = \langle\bar{\mathcal{O}}(M)\rangle(1 + \omega_1\bar{a}(\mu) + \omega_2\bar{a}(\mu)^2 + \cdots). \tag{6.55}$$

The coefficients ν_i and ω_i encode the finite renormalization that relates the two schemes to each other. Taking the difference between the logarithm of Eq. (6.6), evaluated in the two different schemes (and assuming that A_n has a different value in each scheme) yields,

$$\ln\left(\frac{\langle\mathcal{O}(M)\rangle}{A_n}\right) - \ln\left(\frac{\langle\bar{\mathcal{O}}(M)\rangle}{\bar{A}_n}\right) = \int_0^{a(M)} \frac{\gamma(x)}{\beta(x)}\,dx - \int_0^{\bar{a}(M)} \frac{\overline{\gamma}(x)}{\beta(x)}\,dx. \tag{6.56}$$

Expanding this expression in powers of $\bar{a}$ then gives,

$$\ln\left(\frac{\bar{A}_n}{A_n}\right) + \omega_1\bar{a} = \bar{a}\left(\frac{d_1 - \bar{d}_1}{b} + \frac{d\nu_1}{b}\right) + \mathcal{O}\left(\bar{a}^2\right). \tag{6.57}$$

We can evaluate this in the $M \to \infty$ limit, and from this we conclude that,

$$\ln\left(\frac{\bar{A}_n}{A_n}\right) = 0, \tag{6.58}$$

and therefore that A_n is scheme invariant. Note that this is subtly different from Λ, which does depend on the subtraction procedure, as can be seen by comparing Eq. (1.113) with Eq. (6.56).

6.2.2 The explicit FRS dependence of r_i

The partial derivatives derived above allow us to obtain the explicit FRS dependence of the coefficients r_i. Integrating Eqs. (6.21) and (6.30), gives,

$$r_1 = \frac{d}{b}\tau_M - \frac{d_1}{b} - X_0(Q), \tag{6.59}$$

where $\tau_M = b\ln\left(M/\tilde{\Lambda}\right)$. $X_0(Q)$ is an FRS invariant quantity, generated as a constant of integration. It is the equivalent of Eq. (1.133) but generalized to the case where we have dependence on *two* arbitrary scales (μ and M). In analogy with the one scale problem, we can define an FRS invariant, non-universal scale parameter, $\Lambda_{\mathcal{M}}$, via $X_0(Q)$. Thus,

$$\frac{d}{b}\tau_M - \frac{d_1}{b} - r_1 = X_0(Q) \equiv d\,\ln\left(\frac{Q}{\Lambda_{\mathcal{M}}}\right). \tag{6.60}$$

In the case of r_2, systematically integrating the partial derivatives in Eqs. (6.22), (6.26), (6.31), (6.35) and (6.48) gives,

$$r_2 = \left(\frac{1}{2} - \frac{b}{2d}\right)r_1^2 + \frac{b}{d}r_1\tilde{r}_1 + \frac{d_1}{d}r_1 - \frac{dc_2}{2b} + \frac{d_1^2}{2bd} + \frac{cd_1}{2b} - \frac{d_2}{2b} + X_2, \tag{6.61}$$

where we have defined,

$$\tilde{r}_1 \equiv r_1(M=\mu) \tag{6.62}$$

$$= \frac{d}{b}\tau_\mu - \frac{d_1}{b} - X_0(Q). \tag{6.63}$$

Again, X_2 is an FRS invariant constant of integration, and is the analogue of the X_2 in Eq. (1.158).

Finally, for r_3 we have,

$$\begin{aligned} r_3 = {} & \frac{c_2dc}{3b} - \frac{c_3d}{6b} + \frac{cd_2}{3b} - \frac{d_3}{3b} - \frac{2d_1^3}{3bd^2} - \frac{cd_1^2}{2bd} - \frac{c^2d_1}{3b} + \frac{d_1d_2}{bd} + \frac{c_2d_1}{3b} \\ & - \frac{2d_1^2r_1}{d^2} - \frac{d_1r_1^2}{d} + \frac{b^2r_1^3}{3d^2} - \frac{r_1^3}{3} - \frac{bcr_1^2}{2d} + \frac{d_2r_1}{d} \\ & - \frac{2bd_1r_1\tilde{r}_1}{d^2} - \frac{b^2r_1\tilde{r}_1^2}{d^2} - \frac{br_1^2\tilde{r}_1}{d} + \frac{bcr_1\tilde{r}_1}{d} \end{aligned}$$

$$+ \quad \frac{2b\tilde{r}_1 r_2}{d} + \frac{2d_1 r_2}{d} + r_1 r_2 + X_3. \tag{6.64}$$

Using Eq. (6.61), we can write this as a multinomial in r_1 and $\tilde{r}_1$, in analogy with Eq. (1.143). We can also derive similar equations for $r_{i>3}$.

References [39,195] contained errors in the equations obtained for the following derivatives: $M\frac{\partial r_3}{\partial M}$, $\frac{\partial r_2}{\partial d_1}$, $\frac{\partial r_3}{\partial d_1}$, $\frac{\partial r_3}{\partial d_2}$, $\frac{\partial r_2}{\partial c_2}$, $\frac{\partial r_3}{\partial c_2}$ and $\frac{\partial r_3}{\partial c_3}$. Consequently, the expressions for r_2, r_3, X_2 and X_3 presented in those works, and also in Ref. [194], are incorrect.

6.3 Theoretical predictions for the moments

We adopt three different approaches to obtaining theoretical predictions of the moments, each of which endorses a different philosophy for dealing with FRS dependence. In CORGI, we do not explicitly set the renormalization scales M and μ. Rather, the relation between these scales is embedded in the FRS invariant X_0. As a result of this, a set of unphysical logarithms (of M and μ) are resummed. These cancel each other out and in so doing, they build the physical $\log Q/\Lambda_{\mathcal{M}}$ dependence automatically (see Eqs. (1.163) - (1.166)). Consequently we can remove the dependence of the result on factorization and renormalization scales. This is a natural generalization of the CORGI approach of section 1.8.3 to the case where we have dependence on two arbitrary scales (M and μ), see Ref. [39].

By contrast, in the physical scale (PS) approach, the scales are explicitly set to Q. The third approach we use lies halfway in between PS and CORGI. We treat the moments as a single effective charge, setting $M = \mu$. We then 'CORGI-ize' this expression, as in the single-scale case, viz section 1.8.3.

6.3.1 CORGI predictions

In analogy with section 1.8.3, we can use the FRS invariant quantities appearing in Eqs. (6.59), (6.61) and (6.64), to obtain an expression in which all RG-predictable terms are resummed. In this way, the physical Q-dependence of $\mathcal{M}(n;Q^2)$ is built automatically via the resummation of unphysical logs [39]. The result of this resummation is simply equivalent to Eq. (6.8) (with the expressions for r_i given in Eqs. (6.61) and (6.64)), evaluated in the 't Hooft scheme. Previously, we stated that the values

of the parameters $\{M, \mu, c_2, \ldots, d_1, d_2, \ldots\}$ can be used to define any particular FRS. However, using Eqs. (6.59) and (6.63) we can swap the first two of these parameters for r_1 and $\tilde{r}_1$. Consequently, the 't Hooft scheme corresponds to the case in which the set of coefficients $\{r_1, \tilde{r}_1, c_2, \ldots, d_1, d_2, \ldots\}$ are all set to zero. Thus, from Eq. (6.8) we have,[1]

$$\mathcal{M}(n;Q^2) \;=\; A_n \left(\frac{ca_0}{1+ca_0}\right)^{d/b} \left(1 + X_2 a_0^2 + X_3 a_0^3 + \ldots\right). \tag{6.65}$$

In this expression, the second term resums an infinite set of terms which are RG-predictable at NLO; specifically, these are all the terms in the full perturbative expansion which do not include factors of X_i. In the third term, each successive $X_i a_0^{i+1}$ component we include resums a further subset of terms which are RG-predictable at $\mathrm{N^iLO}$. As noted above, this has the effect of resumming the RG-predictable set of unphysical logarithms ($\ln M$ and $\ln \mu$), and through Eqs. (6.59) and (6.63), builds the physical Q^2-dependence automatically.

Setting $r_1 = \tilde{r}_1 = d_1 = 0$ also implies $a = \tilde{a} = a_0$, where a_0 is equivalent to the coupling evaluated in the 't Hooft scheme,

$$a_0(Q) \;=\; \frac{-1}{c\left[1 + W_{-1}(z(Q))\right]}, \tag{6.66}$$

with,

$$z(Q) \;=\; -\frac{1}{\mathrm{e}}\left(\frac{Q}{\Lambda_{\mathcal{M}}}\right)^{-b/c}, \tag{6.67}$$

and $W_{-1}(z)$ defined in Eq. (1.116). The scale $\tilde{\Lambda}_{\mathcal{M}}$, can be related to the scale parameter for any particular scheme using Eq. (6.60). This expression is FRS invariant and hence to relate it to the $\overline{\mathrm{MS}}$ scale, we simply evaluate it in this scheme,

$$\Lambda_{\mathcal{M}} \;=\; \Lambda_{\overline{\mathrm{MS}}} \left(\frac{2c}{b}\right)^{-c/b} \exp\left\{\frac{d_1}{db} + \frac{r_1}{d}\right\}, \tag{6.68}$$

with r_1 and d_1 calculated in $\overline{\mathrm{MS}}$ with $M = \mu = Q$.

[1] Note that in the 't Hooft scheme we have $\mathcal{I}(a) = 0$, $r_{i>1} = X_{i>1}$ and $r_1 = 0$.

The FRS invariant X_2 can be obtained from Eq. (6.61),

$$X_2 = r_2 - \left(\frac{1}{2} - \frac{b}{2d}\right) r_1^2 - \frac{b}{d} r_1 \tilde{r}_1 - \frac{d_1}{d} r_1 + \frac{dc_2}{2b} - \frac{d_1^2}{2bd} - \frac{cd_1}{2b} + \frac{d_2}{2b}. \quad (6.69)$$

Again, this can be evaluated in any scheme. In $\overline{\text{MS}}$ with $M = \mu = Q$, we have $r_1 = \tilde{r}_1$ and hence,

$$X_2 = r_2 - \left(\frac{1}{2} + \frac{b}{2d}\right) r_1^2 - \frac{d_1}{d} r_1 + \frac{dc_2}{2b} - \frac{d_1^2}{2bd} - \frac{cd_1}{2b} + \frac{d_2}{2b}\Bigg|_{\overline{\text{MS}}}. \quad (6.70)$$

This can then be evaluated from the $\overline{\text{MS}}$ two-loop coefficients of the anomalous dimension and coefficient functions.

For the $n = 1$ case we have $d = 0$. This would appear to make Eq. (6.65) invalid, due to the $1/d$ terms in Eqs. (6.68) and (6.70). However, using Eq. (6.60) we can show that the coefficients of the $1/d$ terms are themselves proportional to d, and hence $\Lambda_{\mathcal{M}}/\Lambda_{\overline{\text{MS}}}$ and X_2 are in fact non-singular. The question remains however, how may we write the $n = 1$ moment in the CORGI form?

Expanding Eq. (6.8) in powers of a with $d = 0$ yields a result which resembles a single-scale perturbation series. Setting $a = \tilde{a}$ then gives,

$$\mathcal{M}(1; Q^2) = A_n \left(1 + R_1 a + R_2 a^2\right), \quad (6.71)$$

with $R_1 = -1$ for all n. R_2 is obtainable from Eq. (6.8) and is given in Eq. (6.76). CORGI-izing the above expression as in the single-scale case then yields:

$$\mathcal{M}(1; Q^2) = A_n \left(1 - a_0\right), \quad (6.72)$$

where a_0 is the coupling in Eq. (6.66) but with $\Lambda_{\mathcal{M}}$ now given by,

$$\Lambda_{\mathcal{M}} = \Lambda_{\overline{\text{MS}}} \left(\frac{2c}{b}\right)^{-c/b} \exp\left\{-\frac{R_2}{b}\right\}. \quad (6.73)$$

6.3.2 Physical scale predictions

In the standard physical scale approach, we set $Q = M = \mu$ and adopt $\overline{\text{MS}}$. Setting $M = \mu$ implies that $a = \tilde{a}$, and hence the moments have the form,

$$\mathcal{M}(n;Q^2) = A_n \left(\frac{ca}{1+ca}\right)^{d/b} \left(1 + R_1 a + R_2 a^2 + \ldots\right). \tag{6.74}$$

The coefficients R_i can be determined by expanding Eq. (6.8) in powers of a,

$$R_1 = r_1 + \frac{d_1}{b} \tag{6.75}$$

$$R_2 = r_2 + \frac{d_1^2}{2b^2} - \frac{cd_1}{2b} + \frac{r_1 d_1}{b} - \frac{dc_2}{2b} + \frac{d_2}{2b}, \tag{6.76}$$

and the coupling in this expression is that of Eq. (1.118) with scale parameter $\Lambda_{\overline{\text{MS}}}$. For the $n = 1$ case, the PS prediction is still given by Eq. (6.74) but with the $\left(\frac{ca}{1+ca}\right)^{d/b}$ term absent.

6.3.3 Effective charge predictions

We can also 'CORGI-ize' the expression for the moments as a single-scale problem. This is done by explicitly setting $\mu = M$ and treating the resultant expression as an effective charge (see section 1.8.4 and Refs. [41,42]) in which we have only one arbitrary scale. By resumming Eq. (6.74), we can obtain the moments in the form,

$$\mathcal{M}(n;Q^2) = A_n \left(c\tilde{R}(a)\right)^{d/b}, \tag{6.77}$$

where,

$$\tilde{R}(a) = a + \tilde{R}_1 a^2 + \tilde{R}_2 a^3 + \ldots, \tag{6.78}$$

is our effective charge and the coefficients $\tilde{R}_i$ have the form,

$$\begin{aligned} \tilde{R}_1 &= \frac{bR_1}{d} - c \\ &= \frac{bcr_1}{d} + \frac{d_1 c}{d} - c^2, \\ \tilde{R}_2 &= \frac{bR_2}{d} + \frac{b^2 R_1^2}{2d^2} - \frac{bR_1^2}{2d} - \frac{bR_1 c}{d} + c^2 \end{aligned} \tag{6.79}$$

$$= \frac{bcr_2}{d} + \frac{bd_1r_1c}{d^2} + \frac{d_2c}{2d} - \frac{c_2c}{2}$$

$$+ \frac{b^2r_1^2c}{2d^2} - \frac{br_1^2c}{2d} + \frac{d_1^2c}{2d^2} - \frac{br_1c^2}{d} - \frac{3d_1c^2}{2d} + c^3. \tag{6.80}$$

We can then 'CORGI-ize' $\tilde{R}(a)$ such that the moments have the form,

$$\mathcal{M}(n;Q^2) = A_n c^{d/b}\left(a_0 + \tilde{X}_2 a_0^3 + \ldots\right)^{d/b}, \tag{6.81}$$

where the coupling a_0 differs from that in Eq. (6.66). In this case, the $\tilde{X}_i$ coefficients are the single-scale RS invariants of Eqs. (1.158) and (1.159). They have the form,

$$\tilde{X}_2 = \tilde{R}_2 - \tilde{R}_1^2 - c\tilde{R}_1 + c_2, \tag{6.82}$$

$$\tilde{X}_3 = \tilde{R}_3 - 3\tilde{R}_1\tilde{R}_2 + 2\tilde{R}_1^3 + \frac{c\tilde{R}_1^2}{2} - \tilde{R}_1c_2 + \frac{1}{2}c_3. \tag{6.83}$$

The coupling in Eq. (6.81) is that of Eq. (6.66) but with the scale $\Lambda_{\mathcal{M}}^{\text{EC}}$ now defined by,

$$\Lambda_{\mathcal{M}}^{\text{EC}} = \left(\frac{2c}{b}\right)^{-c/b} \exp\left(\frac{\tilde{R}_1}{b}\right) \Lambda_{\overline{\text{MS}}}, \tag{6.84}$$

(see Eq. (1.134)). For the $n = 1$ case, the EC prediction for the moments is identical to the CORGI prediction, given in Eq. (6.72)

The coefficients of the three predictions in Eqs. (6.65), (6.72), (6.74) and (6.81) are all built from the coefficients for the anomalous dimension and of the coefficient function, which are obtainable (up to NNLO) from Refs. [78] and [113] respectively. In appendix C can be found the values for the NNLO coefficients X_2, R_2 and $\tilde{X}_2$, the NLO coefficient R_1, and the NLO information relevant to the CORGI and effective charge approaches i.e. $\Lambda_{\mathcal{M}}/\Lambda_{\overline{\text{MS}}}$ and $\Lambda_{\mathcal{M}}^{\text{EC}}/\Lambda_{\overline{\text{MS}}}$. We state their values for $N_f = 4$ and $N_f = 5$, and for the range $1 \leq n \leq 20$.

6.3.4 Target mass corrections and higher twist contributions

The perturbative expressions derived above will be subject to non-perturbative corrections in the form of $\mathcal{O}\left(1/Q^2\right)$ terms. The two principal sources of these terms are: higher twist terms (of the type neglected in the OPE analysis of section 2.4) and effects due to the mass of the target hadron.

Target mass corrections

The perturbative form of the moments is derived under the assumption that the mass of the target hadron is zero (in the limit $Q^2 \to \infty$). At intermediate and low Q^2 this assumption will begin to break down and the moments will be subject to potentially significant power corrections of order $\mathcal{O}\left(m_N^2/Q^2\right)$, where m_N is the nucleon mass. These are known as target mass corrections (TMCs) and when they are included, the F_3 moments have the form [198, 199],

$$\mathcal{M}^{\mathrm{TMC}}(n;Q^2) = \mathcal{M}(n;Q^2) + \frac{n(n+1)}{n+2}\frac{m_N^2}{Q^2}\mathcal{M}(n+2;Q^2) + \mathcal{O}\left(\frac{m_N^4}{Q^4}\right). \quad (6.85)$$

Higher twist corrections

The moments will also be subject to corrections from sub-leading twist contributions to the OPE. These effects are poorly understood and hence we only estimate them. This is done by means of an unknown parameter, A_{HT}. The estimate has the form [192],

$$\mathcal{M}^{\mathrm{HT}}(n;Q^2) = n\left(A^{\mathrm{HT}}\frac{\Lambda^2_{\overline{\mathrm{MS}}}}{Q^2}\right)\mathcal{M}(n;Q^2), \quad (6.86)$$

and the value of A^{HT} is obtained by fitting to data. Due to the poorly understood nature of these effects, we do not include the above term in the full analysis. Rather, we perform the analysis with and without this term included, and take the difference in the results as an estimate of the error associated with our ignorance of the true nature of these effects.

6.3.5 Quark mass thresholds

The bottom quark mass threshold is within the range of Q^2 that is spanned by the available data for F_3. It is therefore necessary to evolve the expressions for the moments over this threshold and for this, we adopt the approach previously detailed in section 5.3. We use massless QCD with 4 quarks ($N_f = 4$) for $Q^2 \leq m_b^2$ and massless QCD with 5 quarks ($N_f = 5$) for $Q^2 > m_b^2$. Here, m_b is the pole mass of the b-quark, and $m_b = 4.85 \pm 0.15$ GeV [32]. We also perform matching of the coupling, resulting in the use of different values of the scale parameter above and below the threshold, denoted $\Lambda^{(5)}_{\overline{\mathrm{MS}}}$ and $\Lambda^{(4)}_{\overline{\mathrm{MS}}}$ respectively. The relation between $\Lambda^{(5)}_{\overline{\mathrm{MS}}}$ and $\Lambda^{(4)}_{\overline{\mathrm{MS}}}$ is governed by the expressions in Eqs. (5.24) - (5.26).

Further to this, we also demand continuity of the moments at the threshold, i.e.

$$\mathcal{M}(n;m_b^2)\big|_{N_f=4} = \mathcal{M}(n;m_b^2)\big|_{N_f=5}. \tag{6.87}$$

As a consequence of this, the parameters A_n also have different values in the $N_f = 4$ and $N_f = 5$ regions and their values are related by,

$$A_n^{(5)} = \left(\frac{A_n}{\mathcal{M}(n;m_b^2)}\bigg|_{N_f=5}\right)\left(\frac{\mathcal{M}(n;m_b^2)}{A_n}\bigg|_{N_f=4}\right)A_n^{(4)}. \tag{6.88}$$

6.4 An alternative derivation of $X_i(n)$

We can derive the FRS invariants X_i via a simpler method to that used in section 6.2. If we set $M = \mu$ in Eq. (6.8), then the moments reduce to a single-scale problem viz section 1.8.3. We can then rearrange the resultant expression in terms of an effective charge, $\hat{R}(a)$,

$$\mathcal{M} = A_n\left(\frac{c\hat{R}(a)}{1+c\hat{R}(a)}\right)^{d/b} \tag{6.89}$$

$\hat{R}(a)$ has the form,

$$\hat{R}(a) = a + \hat{R}_1 a^2 + \hat{R}_2 a^3 + \ldots. \tag{6.90}$$

The coefficients $\hat{R}_i$ can be determined by expanding Eqs. (6.8) and (6.89) in powers of a and then equating coefficients. They are found to be,

$$\hat{R}_1 = \frac{b}{d}R_1, \tag{6.91}$$

$$\hat{R}_2 = \frac{b}{d}\left(R_2 + cR_1 - \frac{R_1^2}{2} + \frac{bR_1^2}{2d}\right), \tag{6.92}$$

where R_1 and R_2 are given by Eqs. (6.75) and (6.76). If we then CORGI-ize this effective charge, we have (from Eqs. (1.133) and (1.158)) a new set of FRS invariants,

$$\hat{X}_0 = b\ln\frac{M}{\tilde{\Lambda}} - \hat{R}_1, \tag{6.93}$$

$$\hat{X}_2 = \hat{R}_2 - \hat{R}_1^2 - c\hat{R}_1 + c_2 \tag{6.94}$$

and the moments become,

$$\mathcal{M} = A_n \left(\frac{c\left(a_0 + \hat{X}_2 a_0^3\right)}{1 + c\left(a_0 + \hat{X}_2 a_0^3\right)} \right)^{d/b} . \tag{6.95}$$

Expanding this into a form which we can compare with Eq. (6.65), gives,

$$\mathcal{M} = A_n \left(\frac{ca_0}{1 + ca_0} \right)^{d/b} \left(1 + \frac{d}{b} \hat{X}_2 \, a_0^2 + \ldots \right). \tag{6.96}$$

Isolating the $\mathcal{O}(a_0^2)$ term in the right hand bracket of the above equation, and then using Eqs. (6.94), (6.91), (6.92), (6.75) and (6.76), gives,

$$\frac{d}{b}\hat{X}_2 = \frac{d}{b}\left(\hat{R}_2 - \hat{R}_1^2 - c\hat{R}_1 + c_2 \right) \tag{6.97}$$

$$= r_2 - \left(\frac{1}{2} + \frac{b}{2d} \right) r_1^2 - \frac{d_1}{d} r_1 + \frac{dc_2}{2b} - \frac{d_1^2}{2bd} - \frac{cd_1}{2b} + \frac{d_2}{2b}. \tag{6.98}$$

So we see that the coefficient of the $\mathcal{O}(a^2)$ term in Eq. (6.96) is the FRS invariant X_2, of Eq. (6.69) with $\mu = M$ i.e. $r_1 = \tilde{r}_1$. Isolating the a^3 term will yield X_3, and so on for higher X_i.

The coupling in Eq. (6.96) is the 't Hooft coupling of Eq. (6.66), but with the scale parameter determined by Eq. (6.93). Evaluation of Eq. (6.93) in $\overline{\text{MS}}$ with $M = \mu = Q$ gives,

$$\hat{X}_0 \equiv b \ln \frac{Q}{\Lambda_{\mathcal{M}}} \tag{6.99}$$

$$= b \ln \frac{Q}{\tilde{\Lambda}} - \hat{R}_1. \tag{6.100}$$

Using Eqs. (6.91) and (6.75) we can then relate $\Lambda_{\mathcal{M}}$ to $\Lambda_{\overline{\text{MS}}}$,

$$\Lambda_{\mathcal{M}} = \Lambda_{\overline{\text{MS}}} e^{\hat{R}_1/b} \tag{6.101}$$

$$= \tilde{\Lambda}_{\overline{\text{MS}}} \exp\left\{ \frac{d_1}{db} + \frac{r_1}{d} \right\}, \tag{6.102}$$

which is in agreement with Eq. (6.68).

This gives us a means by which we can obtain the CORGI coefficients without having to

calculate the partial derivatives of section 6.2.1. Moreover, we can now obtain CORGI predictions from expressions in which μ has already been set to M. This will be useful when we come to consider singlet components of moments of F_2 in the next chapter.

6.5 The method of Bernstein averages

When comparing theoretical predictions for moments of structure functions with experimental data, we are faced with the long-standing issue of missing data regions at high and low x for low and high Q^2 respectively. This is demonstrated in Fig. 6.1 in which we plot the Columbia-Chicago-Fermilab-Rochester (CCFR) data [200] for 12 different values of Q^2. We can see that at the lower range of Q^2 we are limited to low-x data, and that at high-Q^2 we are limited to the high-x range.

In order to reliably evaluate a moment at a particular Q^2, we require data for the whole range of x. This being unavailable, we are forced to make some guess about how the structure function behaves in the missing data region. That is to say, we have to choose some method of modelling (extrapolating and interpolating) the data to cover the full range of x. We could use some QCD inspired method for this modelling, such as MRST [201] or CTEQ [202]. However, we wish to keep our experimental input as free from QCD assumptions as possible – the intent being to make the comparison of theoretical predictions for the moments with experiment, as direct as possible. Therefore we carry out this comparison indirectly, via Bernstein averages – objects which, though related to the moments, have negligible dependence on the modelling method adopted (and hence on the behaviour of the structure function in the missing data regions).

We define the Bernstein polynomials as follows [203],

$$p_{nk}(x^2) = 2\frac{\Gamma\left(n+\frac{3}{2}\right)}{\Gamma\left(k+\frac{1}{2}\right)\Gamma\left(n-k+1\right)}x^{2k}(1-x^2)^{n-k}, \qquad n,k \in \mathbb{I}. \quad (6.103)$$

These functions are constructed such that they are zero at the endpoints $x = 0$ and $x = 1$, and they are also normalized such that $\int_0^1 p_{nk}(x)dx = 1$. Furthermore, if we constrain n and k such that $n > k \geq 0$, then $p_{nk}(x)$ are peaked sharply in some region between the two endpoints.

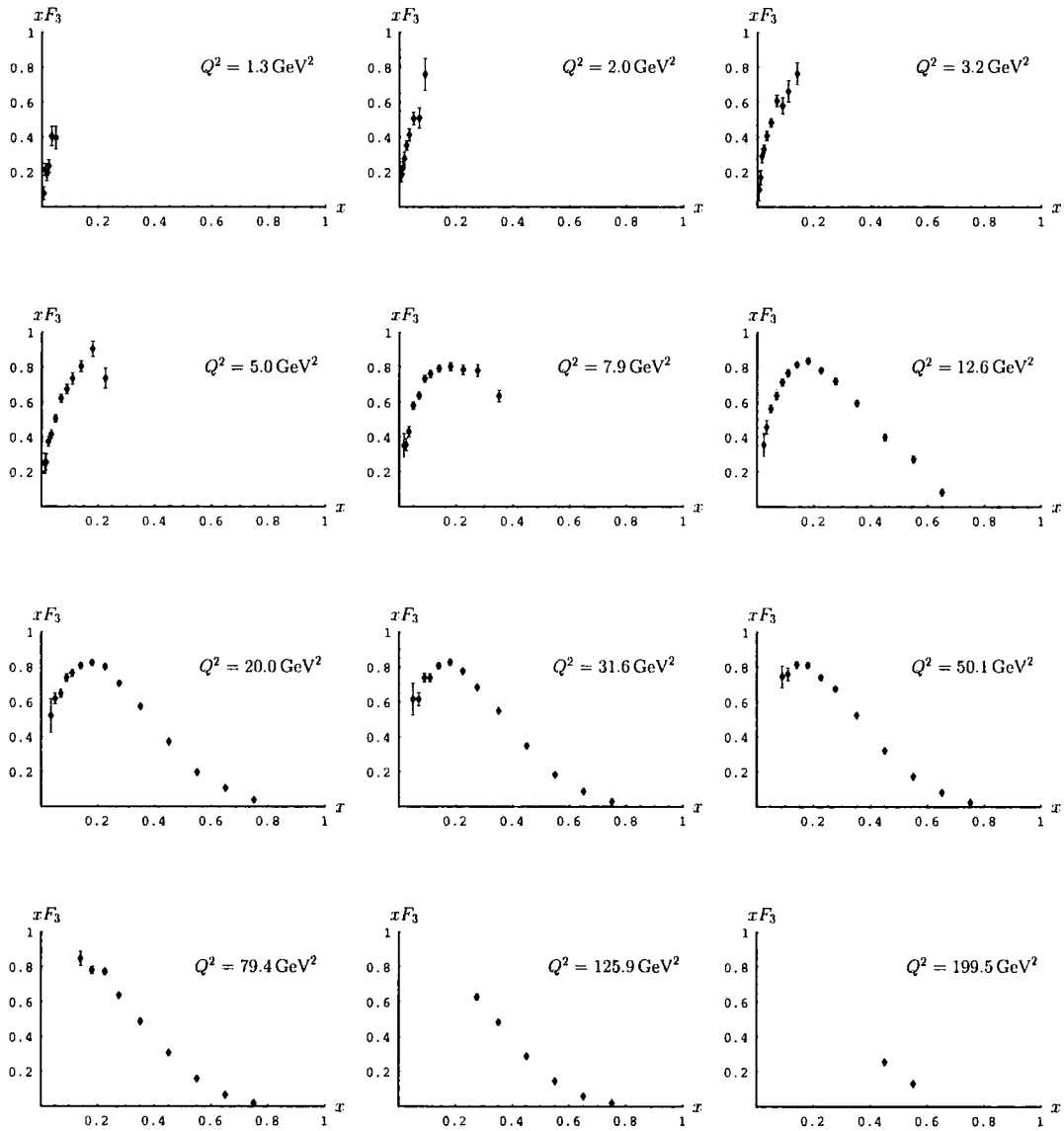


Figure 6.1: Data for xF_3 plotted against x for the 12 different Q^2 bins of the CCFR data [200].

The Bernstein polynomials can be treated as a distribution, with a mean,

$$\overline{x}_{nk} \equiv \int_0^1 x\, p_{nk}(x)\, dx \tag{6.104}$$

$$= \frac{\Gamma(k+1)\Gamma(n+\frac{3}{2})}{\Gamma(k+\frac{1}{2})\Gamma(n+2)}, \tag{6.105}$$

and variance,

$$\Delta x_{nk} \equiv \int_0^1 (x-\overline{x}_{nk})^2\, p_{nk}(x)\, dx \tag{6.106}$$

$$= \frac{k+\frac{1}{2}}{n+\frac{3}{2}} - \left(\frac{\Gamma(k+1)\Gamma(n+\frac{3}{2})}{\Gamma(k+\frac{1}{2})\Gamma(n+2)}\right)^2. \tag{6.107}$$

The Bernstein averages of F_3 are then defined by,

$$F_{nk}(Q^2) = \int_0^1 p_{nk}(x^2) F_3(x, Q^2) dx. \tag{6.108}$$

Thus F_{nk} is the average of the structure function weighted such that the region around $\overline{x}_{nk}$ is emphasized. By picking the values of n and k wisely, we can construct a set of averages which enhance the region for which we have data for F_3, and de-emphasize the regions where there are gaps. Therefore, in the resultant averages, the dependence on the missing data regions will be heavily suppressed.

Defining this more rigorously: for a given value of Q^2, we only consider averages for which the range,

$$\overline{x}_{nk} - \frac{1}{2}\sqrt{\Delta x_{nk}} \leq x \leq \overline{x}_{nk} + \frac{1}{2}\sqrt{\Delta x_{nk}}, \tag{6.109}$$

lies entirely within the region for which we have data. The only exception to this is if the highest-x data point lies within this range, then we *do* accept this average, but only if the data suggests that xF_3 vanishes very rapidly beyond this point.

The construction of an acceptable average, and the resultant suppression of the missing data region is demonstrated in Fig. 6.2. We see that the red missing data regions disappear almost completely in the right hand plot.

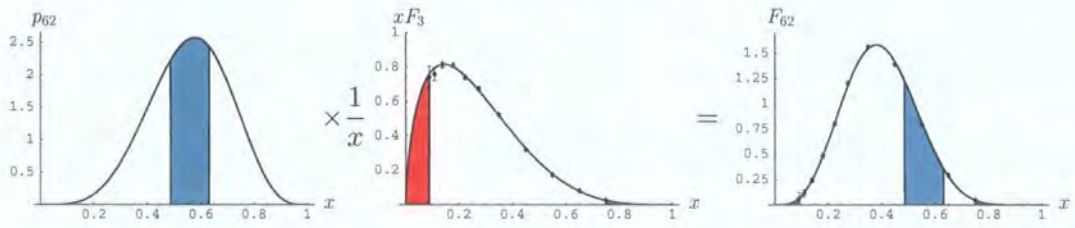


Figure 6.2: Constructing the Bernstein average, $F_{62}(Q^2 = 50.1)$. The cyan region represents the interval in Eq. (6.109) and the red areas represent the missing data regions. The small size of the red region in the right hand plot demonstrates that this average will have negligible dependence on the missing data regions. Note that the right hand plot actually shows the *integrand* of the Bernstein average. The average itself will be this function integrated over $[0, 1]$.

By expanding the integrand of Eq. (6.108) in powers of x, and using Eq. (6.1), we can relate the averages directly to the moments,

$$F_{nk}(Q^2) = \frac{2\Gamma(n+\frac{3}{2})}{\Gamma(k+\frac{1}{2})} \sum_{l=0}^{n-k} \frac{(-1)^l}{l!(n-k-l)!} \mathcal{M}(2(k+l)+1; Q^2), \qquad (6.110)$$

and so theoretical predictions for the averages can be obtained by substitution of Eqs. (6.65), (6.72) (6.74) and (6.81) into the above expression.

Modelling the structure functions

In order to calculate averages from data for F_3, we need an expression for xF_3 covering the entire range of x, for each value of Q^2.

As mentioned above, the values of the moments calculated in this way will depend on how we model the structure functions in the missing data regions, but for the averages this dependence is suppressed. However, we would like to test this assertion, and so we use four different methods of modelling F_3 and perform our analysis separately for each method. Significant differences between the results would signify a failure of the Bernstein average method, and in instances where this is the case, we reject that particular average at that particular Q^2. Moderate deviation, however, is acceptable, provided that we use the magnitude of the deviation as an estimate of the error associated with the missing data region. This error is then included as a 'modelling error' in the final result. In this way, we can almost completely remove any dependence on missing data

regions, and quantify the error associated with any residual dependence.

The four extrapolation methods we use are described below:

I In the first method, we fit the function,

$$xF_3(x) \;=\; A\,x^B(1-x)^C, \tag{6.111}$$

to the data for each fixed value of Q^2. The parameters A, B and C are obtained by performing χ^2 fitting of Eq. (6.111) to data for F_3. They are Q^2 dependent quantities, and errors on their values are obtained by performing the fitting with the data for F_3 shifted to the two extremes of the error bars.

A justification for the particular form of fitting function in Eq. (6.111) can be found in Ref. [204]. However, the simple fact that this function fits the data well is justification enough, since the Bernstein averages are independent of the extrapolation method.

II The second method we use is linear interpolation between successive data points. We also extrapolate beyond the data range, to the endpoints $xF_3(x)|_{x=0} = xF_3(x)|_{x=1} = 0$, in order to be consistent with method **I**.

III The third method consists of using the fitting function of Eq. (6.111), but setting $xF_3(x) = 0$ everywhere outside the region for which we have data.

IV In analogy with **III**, in this method we use the linear interpolation of method **II** but setting $xF_3(x) = 0$ everywhere outside the data region.

The deviation between the results obtained from the above methods (in particular, the difference between the first two and the last two) will be a good measure of the effectiveness of the Bernstein average method.

In Fig. 6.3 we show each of the four modelling methods applied to F_3 measured at $Q^2 = 79.4$ GeV^2. Also shown on these figures (in grey) are the fits which are used to determine the statistical error on the form of F_3, which then propagates through to errors on the averages. The systematic errors are determined in the same way.

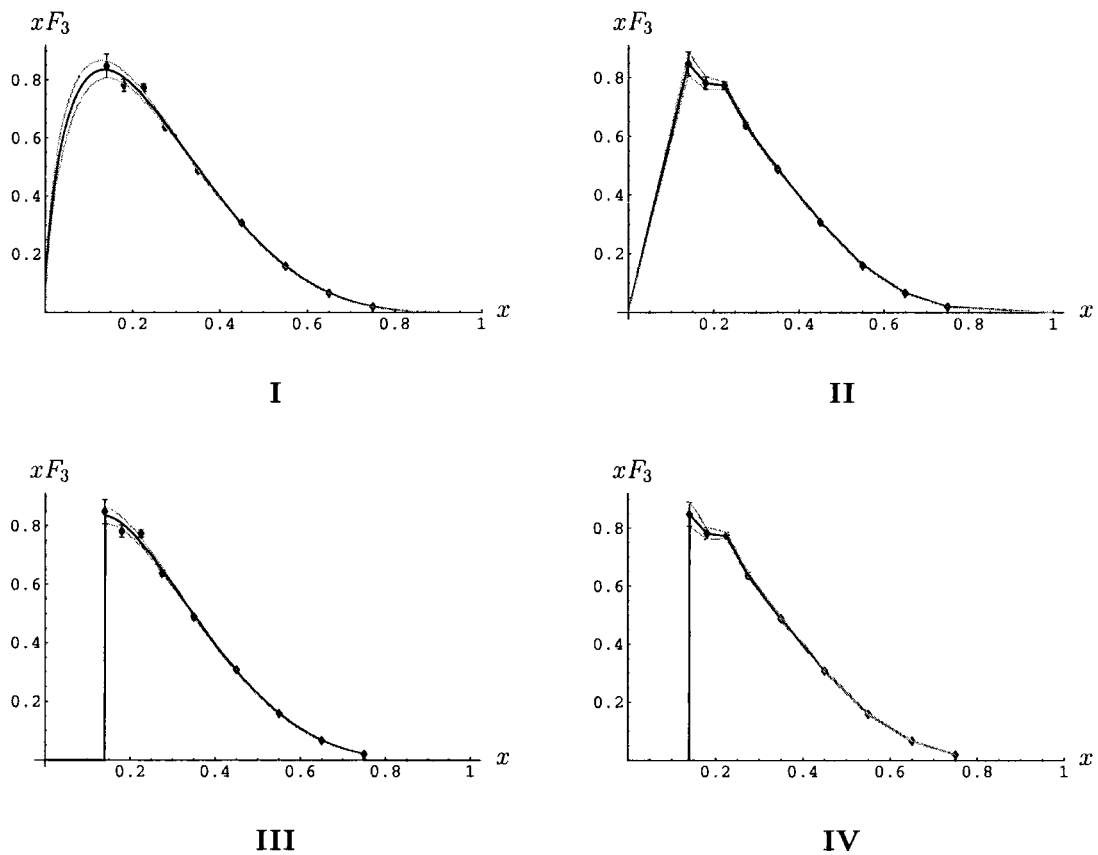


Figure 6.3: The four methods used for modelling the structure functions. Here we show the measured values of xF_3 at $Q^2 = 79.4\,\mathrm{GeV}^2$. The statistical errors are determined by re-performing the fitting for the data shifted to the extremes of the (statistical) error bars, and this is denoted by the grey lines.

6.5.1 Modified Bernstein averages

Due to the unavailability of results for d_2 for even n, previous NNLO analyses of this kind have been limited to the inclusion of only odd F_3 moments. As an artefact of this, the Bernstein polynomials of (6.103) are constructed such that the resultant Bernstein averages depend only on odd moments, as can be seen from Eq. (6.110).

However, now that the NNLO calculation of the NS anomalous dimension is complete [78], we are no longer constrained in such a way. In light of this, we define a new set of *modified* Bernstein polynomials,

$$\tilde{p}_{nk}(x^2) = 2\frac{\Gamma(n+2)}{\Gamma(k+1)\Gamma(n-k+1)}x^{2k+1}(1-x^2)^{n-k}, \qquad n,k \in \mathbb{I}, \tag{6.112}$$

which include only odd powers of x and hence whose averages are related to even moments. These modified Bernstein polynomials are simply the original polynomials of Eq. (6.103), multiplied by x, and then "re-normalized" such that they still satisfy $\int_0^1 \tilde{p}_{nk}(x)dx = 1$. We can calculate the mean and variance of $\tilde{p}_{nk}(x)$,

$$\bar{\tilde{x}} = \frac{\Gamma(k+\frac{3}{2})\Gamma(n+2)}{\Gamma(k+1)\Gamma(n+\frac{5}{2})}, \tag{6.113}$$

$$\Delta\tilde{x} = \frac{k+1}{n+2} - \left[\frac{\Gamma(k+\frac{3}{2})\Gamma(n+2)}{\Gamma(k+1)\Gamma(n+\frac{5}{2})}\right]^2, \tag{6.114}$$

and in analogy with Eq. (6.108), we define the modified Bernstein averages,

$$\tilde{F}_{nk}(Q^2) = \int_0^1 \tilde{p}_{nk}(x^2)F_3(x,Q^2)dx. \tag{6.115}$$

Again, we only accept experimental modified averages for which the range,

$$\bar{\tilde{x}}_{nk} - \frac{1}{2}\sqrt{\Delta\tilde{x}_{nk}} \leq x \leq \bar{\tilde{x}}_{nk} + \frac{1}{2}\sqrt{\Delta\tilde{x}_{nk}}, \tag{6.116}$$

lies within the region for which we have data, with the same exception as that we applied to the standard averages. We obtain theoretical predictions for the modified averages using the equation,

$$\tilde{F}_{nk}(Q^2) = \frac{2\Gamma(n+2)}{\Gamma(k+1)}\sum_{l=0}^{n-k}\frac{(-1)^l}{l!(n-k-l)!}\mathcal{M}(2(k+l)+2;Q^2), \tag{6.117}$$

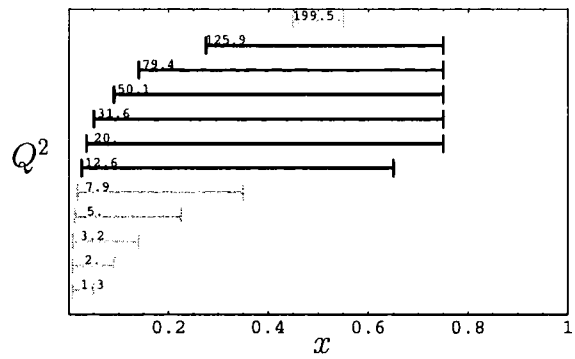


Figure 6.4: Diagram depicting the x-ranges covered by the CCFR data, at different Q^2.

and the expressions in Eqs. (6.65), (6.72) (6.74) and (6.81).

6.5.2 Experimental input

Data for xF_3 in neutrino-nucleon scattering is available from the CCFR collaboration [200]. The data was obtained from the scattering of neutrinos off iron nuclei and the measurements span the ranges 1.26 GeV$^2 \leq Q^2 \leq$ 199.5 GeV^2 and $0.015 \leq x \leq 0.75$. The x-ranges covered at each Q^2 are depicted in Fig. 6.4.

From this data (and using the methods **I** - **IV** outlined above) we can obtain expressions describing the behaviour of the structure function over the full range of x, for each value of Q^2. It is then possible to extract experimental values of the averages, using the methods outlined below:

In the case of method **I**, obtaining the averages is particularly simple. Substituting Eq. (6.111) into Eqs. (6.108) and (6.115) gives,

$$F_{nk}^{(\mathrm{exp})} = A\frac{2\Gamma(n+\frac{3}{2})}{\Gamma(k+\frac{1}{2})}\sum_{l=0}^{n-k}\frac{(-1)^l}{l!(n-k-l)!}\mathrm{Beta}(2(k+l)+B, C+1), \quad (6.118)$$

for the Bernstein averages and,

$$\tilde{F}_{nk}^{(\mathrm{exp})} = A\frac{2\Gamma(n+2)}{\Gamma(k+1)}\sum_{l=0}^{n-k}\frac{(-1)^l}{l!(n-k-l)!}\mathrm{Beta}(2(k+l)+B+1,C+1), \quad (6.119)$$

for the modified Bernstein averages. Here, $\mathrm{Beta}(x,y)$ is the beta function, defined in Eq. (B.5). Once values for A, B, and C have been obtained, substitution into the above expressions leads directly to the averages.

In the case of method **II**, each of the averages is split into $j+1$ sections (where j is the number of data points at a particular Q^2), and each section is an integral of a polynomial of order $2n+1$. It is then reasonably simple to evaluate the averages by computing this set of integrals. This approach also applies to method **IV**, but in this case there are only $j-1$ integrals.

For method **III** we simply integrate the fitting function, multiplied by the Bernstein polynomials, with the integration limits being the values of x at the first and last data points.

Choosing appropriate averages

Having outlined the method for obtaining the experimental averages we now turn our attention to which averages are acceptable at which energies. The highest moment we use is the 18th moment and the lowest the 1st. Inclusion of higher moments than this leads to *no* significant increase in the number of acceptable Bernstein averages and hence will lead to no improvement on the precision with which we can predict $\Lambda_{\overline{\mathrm{MS}}}$.

The upper limit of $n=18$ implies that the highest Bernstein averages included are F_{8k} and $\tilde{F}_{8k}$ and that the lowest used are F_{10} and $\tilde{F}_{10}$. We exclude the averages for which $n=k$ as they simply correspond to individual moments themselves. This leaves us with a total of 72 (36 + 36) potential averages at our disposal for each value of Q^2. This number will be reduced when we come to exclude averages on the basis of the acceptance criteria.

In Fig. 6.5 we plot the dominant regions of the Bernstein polynomials (given by Eq. (6.109)) for each of the used averages. This is superimposed onto the data-range

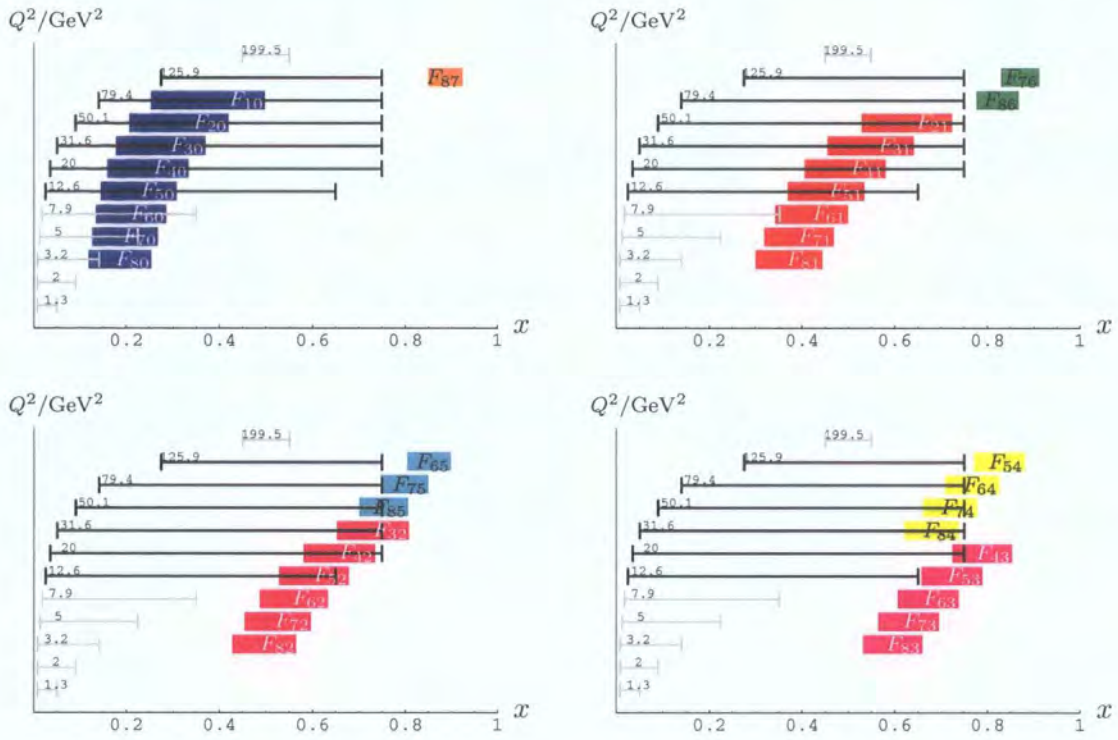


Figure 6.5: The black and light grey bars (⊢⊣) show x ranges covered by the CCFR data at different energies. Superimposed onto these, in various colours, are the peaked regions of the individual Bernstein polynomials, defined by the interval in Eq. (6.109).

diagram of Fig. 6.4. Figure 6.6 shows equivalent plots for the modified Bernstein averages. These plots can be used to identify which averages are acceptable for a particular value of Q^2.

After applying the acceptance criteria, we are left with 132 data points for the standard Bernstein averages and 141 for the modified Bernstein averages. Exactly which averages we use at a particular Q^2, can be determined by inspecting the plots in the results section.

6.6 Fitting procedure

We use χ^2 minimization to optimize the fits of the theoretical predictions to the data. The highest moment included in the experimental averages is the 18th, and so when TMCs are included, we will require predictions for the first 20 moments. Therefore,

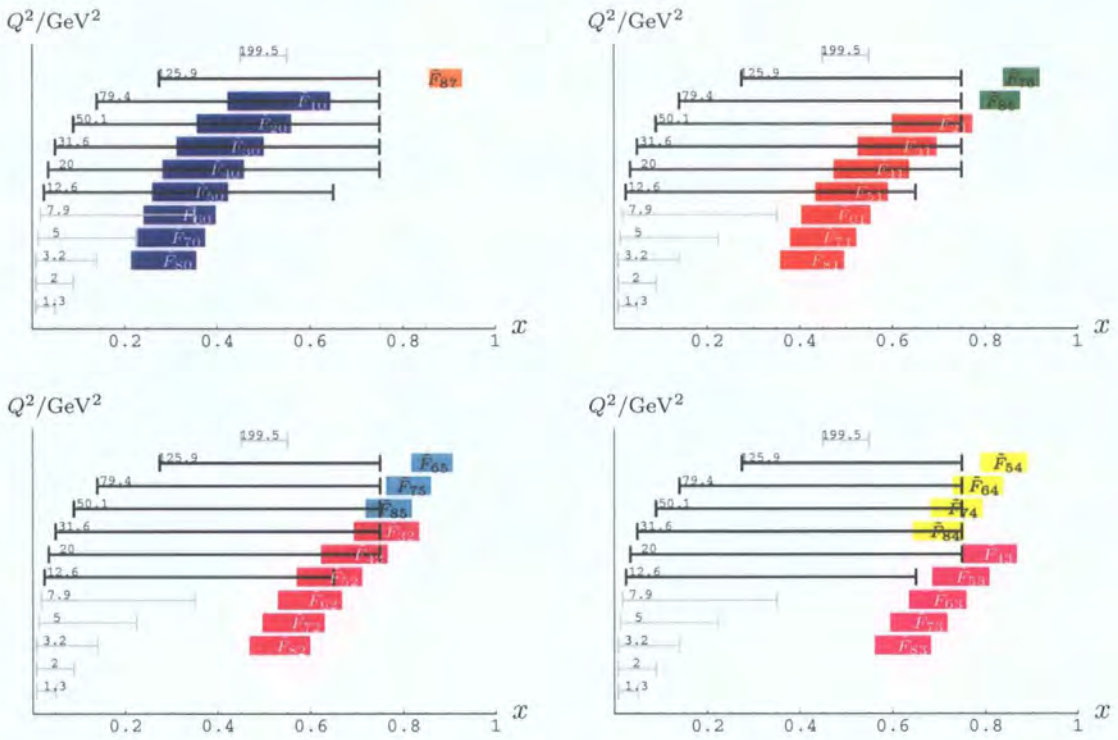


Figure 6.6: The black and light grey bars (⊢⊣) show x ranges covered by the CCFR data at different energies. Superimposed onto these, in various colours, are the peaked regions of the individual modified Bernstein polynomials, defined by the interval in Eq. (6.116).

the set of fitting parameters comprises of $\{A_1 \ldots A_{20}\}$ plus the QCD scale parameter $\Lambda_{\overline{\rm MS}}$. When we include higher twist corrections this set is expanded to include $A_{\rm HT}$. To check consistency between the odd and even moments we perform the analysis for each of these sets of moments separately and then finally together, and compare the results.

Although we stated previously that 132 standard Bernstein averages are available to us, in the CORGI case this is reduced to 130 for the following reason: When fitting predictions to the data, we scan values of $\Lambda^{(4)}_{\overline{\rm MS}}$ between 0 and 590 MeV for a minimum in χ^2. Unfortunately, for $n = 17$ and 19, the values of $\Lambda_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$ (see section C.1) are such that $Q^2 = 7.9{\rm GeV}^2$ is below the Landau pole in Eq. (6.66). Consequently we cannot obtain CORGI predictions for Bernstein averages which include these moments. From Eqs. (6.85) and (6.110) we can determine that this excludes $F_{80}(Q^2 = 7.9)$ and $F_{70}(Q^2 = 7.9)$ from the fit. In the case of the modified Bernstein averages, $\tilde{F}_{80}(Q^2 = 7.9)$ and $\tilde{F}_{70}(Q^2 = 7.9)$ are already excluded due to their failure to meet the acceptance

criteria.

The CCFR data includes statistical errors and 18 different sources of systematic error. These errors cannot be added in quadrature, and so we perform the analysis for each of these 19 sources of error separately and then add the variation in the results in quadrature to obtain the final total error on the Bernstein averages. We also include, as additional sources of error: the deviation in results associated with using the four different modelling methods (this forms the 'modelling error' in our final result) and the deviation in the results obtained by performing the analysis with and without HT corrections included (forming the 'HT error').

Determination of errors

The optimum value of $\Lambda_{\overline{\rm MS}}$ is obtained by minimizing the following χ^2 function with respect to all parameters,

$$\chi^2\left(\{A_n\}, \Lambda_{\overline{\rm MS}}, A_{\rm HT}\right) = \sum_{\substack{n=1 \\ n>k}}^{8}\sum_{k=0}^{8}\sum_{Q^2\in\{Q_i^2\}}\left(\frac{F_{nk}^{\rm theory}(Q^2) - F_{nk}^{\rm exp}(Q^2)}{\sigma F_{nk}^{\rm exp}(Q^2)}\right)^2 + \sum_{\substack{n=1 \\ n>k}}^{8}\sum_{k=0}^{8}\sum_{Q^2\in\{Q_i^2\}}\left(\frac{\tilde{F}_{nk}^{\rm theory}(Q^2) - \tilde{F}_{nk}^{\rm exp}(Q^2)}{\sigma \tilde{F}_{nk}^{\rm exp}(Q^2)}\right)^2. \tag{6.120}$$

Here, $\{Q_i^2\}$ is the set of values of Q^2 for which we have experimental values for the averages. $\sigma F_{nk}^{\rm exp}(Q^2)$ represents the experimental error on the averages and is obtained by adding the statistical, systematic and modelling errors in quadrature:

$$\left(\sigma F_{nk}^{\rm exp}\right)^2 = \left(\sigma F_{nk}^{\rm exp}\right)^2_{\rm stat} + \left(\sigma F_{nk}^{\rm exp}\right)^2_{\rm sys} + \left(\sigma F_{nk}^{\rm exp}\right)^2_{\rm mod} \tag{6.121}$$

We isolate the individual contribution of these three sources of error to the value of the total error on $\Lambda_{\overline{\rm MS}}$ using the following method: For the statistical error, we set $\left(\sigma F_{nk}^{\rm exp}\right)^2 = \left(\sigma F_{nk}^{\rm exp}\right)^2_{\rm stat}$ and then minimize χ^2 with respect to the fitting parameters, to obtain the optimal value of $\Lambda_{\overline{\rm MS}}$ relative to the statistical errors. We then search for the maximum amount $\Lambda_{\overline{\rm MS}}$ can deviate by (in both the positive and negative directions), subject to the constraint that $\chi^2 = \chi^2_{\rm min} + 1$. This yields a value for $(\sigma\Lambda_{\overline{\rm MS}})_{\rm stat}$. We

then repeat this procedure for the systematic and modelling errors to obtain values for $(\sigma\Lambda_{\overline{\rm MS}})_{\rm sys}$ and $(\sigma\Lambda_{\overline{\rm MS}})_{\rm mod}$. The total error on $\Lambda_{\overline{\rm MS}}$ is then obtained by adding these values in quadrature.

Correlation of errors

When fitting theoretical predictions to experimental data using χ^2 minimization, care must be taken in order to take into account fully the correlation between data points.

The raw data for xF_3 is uncorrelated. However, we are not comparing predictions for the structure functions themselves with data directly; rather we are doing so indirectly via the Bernstein averages. For a given value of Q^2, the full set of Bernstein averages (and modified Bernstein averages) we obtain *will* be correlated, due to their being derived from the same set of (xF_3) data points (see Fig. 6.2). Furthermore, we can expect this correlation to be significant, since for a given value of Q^2 the averages which satisfy the acceptance criteria are all derived from Bernstein polynomials that are peaked within a fixed region in x.

In the case where data are correlated, the form of χ^2 given in Eq. 6.120 is not sufficient. In order to take into account the correlation, a more general form of χ^2 involving the covariance matrix $\mathbf{V}$ must be used (see Ref. [205] for details).

Previously, to construct χ^2 from a set of N uncorrelated data points $\{f_i^{\rm exp.}\}$ $(i = 1,\ldots N)$, with errors $\{\sigma_{f_i}\}$ and corresponding theoretical predictions $\{f_i^{\rm theo.}\}$, we had,

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{f_i^{\rm exp.} - f_i^{\rm theo.}}{\sigma_{f_i}} \right)^2 . \tag{6.122}$$

However, for the case where $\{f_i^{\rm exp.}\}$ are correlated we have,

$$\begin{aligned} \chi^2 &= \sum_{i=1}^{N}\sum_{j=1}^{N} (f_i^{\rm exp.} - f_i^{\rm theo.})\, V_{ij}^{-1} \left(f_j^{\rm exp.} - f_j^{\rm theo.} \right) \\ &= (\mathbf{f}^{\rm exp.} - \mathbf{f}^{\rm theo.})^T\, \mathbf{V}^{-1}\, (\mathbf{f}^{\rm exp.} - \mathbf{f}^{\rm theo.}) . \end{aligned} \tag{6.123}$$

In the second line of the above equation we have constructed vectors from the data points and their predictions.

The covariance matrix encodes the correlation between each of the data points; its elements are obtained as follows,

$$\begin{aligned} V_{ij} &= \mathrm{cov}\,(f_i, f_j) \\ &= \langle f_i f_j \rangle - \langle f_i \rangle \langle f_j \rangle. \end{aligned} \tag{6.124}$$

However, in the case where the f_i are functions of M variables x_k (representing the 'raw' data), we have,

$$\mathrm{cov}\,(f_k, f_l) = \sum_{i=1}^{M} \sum_{j=1}^{M} \left(\frac{\partial f_k}{\partial x_i}\right) \left(\frac{\partial f_l}{\partial x_j}\right) \mathrm{cov}\,(x_i, x_j)\,. \tag{6.125}$$

If the data for x_i are uncorrelated, this reduces to,

$$\mathrm{cov}\,(f_k, f_l) = \sum_{i=1}^{M} \left(\frac{\partial f_k}{\partial x_i}\right) \left(\frac{\partial f_l}{\partial x_i}\right) \sigma_{x_i}^2. \tag{6.126}$$

To apply this to data for the Bernstein averages, we first form vectors out of the data. The averages are uncorrelated between different values of Q^2, therefore we can separate the data out into seven separate vectors; one for each value of Q^2 between 7.9 and 125.9 GeV^2. For example, for $Q^2 = 7.9$ and 12.6 GeV^2 we have,

$$\mathbf{F}_{7.9}^{\mathrm{exp.}} = \begin{pmatrix} F_{60}(7.9) \\ F_{50}(7.9) \\ F_{40}(7.9) \end{pmatrix} \qquad \mathbf{F}_{12.6}^{\mathrm{exp.}} = \begin{pmatrix} F_{80}(12.6) \\ F_{70}(12.6) \\ F_{60}(12.6) \\ \vdots \\ F_{52}(12.6) \\ F_{83}(12.6) \\ F_{73}(12.6) \end{pmatrix}. \tag{6.127}$$

We then construct similar vectors, $\mathbf{F}_{Q_i}^{\mathrm{theo.}}$, from the theoretical predictions for the averages.

Finally, the elements of the covariance matrix must be determined. Like the vectors, there are seven separate matrices, one for each value of Q^2. To obtain the elements of these matrices, it is necessary to approximate the averages (using the trapezium rule)

in the following form,

$$F_{nk}(Q_i^2) = \sum_{j=0}^{N_i} \frac{1}{2} \left(p_{nk}(x_j) F_3(x_j; Q_i^2) + p_{nk}(x_{j+1}) F_3(x_{j+1}; Q_i^2) \right) (x_{j+1} - x_j)$$

$$= \sum_{j=1}^{N_i} \frac{1}{2} p_{nk}(x_j) F_3(x_j; Q_i^2) (x_{j+1} - x_{j-1}), \tag{6.128}$$

where N_i is the number of xF_3 data points available for $Q^2 = Q_i^2$. The $j = 0$ and $j = N_i + 1$ terms are simply the $x = 0$ and $x = 1$ endpoints (see parts **I** and **II** of Fig. 6.3). To obtain elements of the covariance matrices, we require the following derivatives,

$$\frac{\partial (F_{nk})_l}{\partial (F_3(x_j))} = \frac{1}{2} (p_{nk}(x_j))_l (x_{j+1} - x_{j-1}). \tag{6.129}$$

This is the differential of the lth element of the vectors in Eq. (6.127) with respect to the jth data point for xF_3. The elements of the covariance matrices can then be obtained as follows,

$$V_{lm}(Q_i^2) = \sum_{j=1}^{N_i} \frac{1}{4} (p_{nk}(x_j))_l (p_{nk}(x_j))_m (x_{j+1} - x_{j-1})^2 \sigma^2_{F_3(x_j)}. \tag{6.130}$$

Hence, the $\mathbf{V}(Q_i^2)$ matrices can be obtained for each value of Q^2 directly from the data for xF_3. Using the same method we can also obtain the covariance matrices for the modified Bernstein averages. In this case there are only 6 matrices, due to there being no acceptable modified Bernstein averages for $Q^2 = 7.9$ GeV^2.

With the covariance matrices included, χ^2 becomes,

$$\chi^2 (\{A_n\}, \Lambda_{\overline{\rm MS}}, A_{\rm HT}) =$$

$$\sum_{Q^2 \in \{Q_i^2\}} \left(\mathbf{F}_{nk}^{\rm theo.}(Q^2) - \mathbf{F}_{nk}^{\rm exp.}(Q^2) \right)^T \mathbf{V}^{-1}(Q^2) \left(\mathbf{F}_{nk}^{\rm theo.}(Q^2) - \mathbf{F}_{nk}^{\rm exp.}(Q^2) \right)$$

$$+ \sum_{Q^2 \in \{Q_i^2\}} \left(\tilde{\mathbf{F}}_{nk}^{\rm theo.}(Q^2) - \tilde{\mathbf{F}}_{nk}^{\rm exp.}(Q^2) \right)^T \tilde{\mathbf{V}}^{-1}(Q^2) \left(\tilde{\mathbf{F}}_{nk}^{\rm theo.}(Q^2) - \tilde{\mathbf{F}}_{nk}^{\rm exp.}(Q^2) \right). \tag{6.131}$$

All that remains is to invert the covariance matrices. However, upon attempting to do

so, we find that these matrices are ill-conditioned, with some of their eigenvalues being close to zero. Hence their inverses are intractable. As a result of this, it is impossible to perform a reliable χ^2 analysis of the averages with their correlation taken into account.

We believe that this is mainly due to the fact that the correlation between averages is significant in some cases, and this in turn is an artefact of the fact that the selection criteria systematically select Bernstein polynomials which are peaked in the same region and hence are of fairly similar shape. This situation arises because the intent behind the inclusion of more averages in the analysis is not to increase the amount of 'data', rather it is to further ensure that the missing data regions are suppressed.

In light of this, we settle for the method adopted in Refs. [192, 193], in which the naïve χ^2 function of Eq. (6.120) is used, but the error bars on the averages are modified in order to account for the 'over-counting of degrees of freedom'.

For example, for the standard averages, at each value of Q^2 we have theoretical information on 9 moments, but the number of experimental Bernstein averages we use is often more than this; e.g. for $Q^2 = 20\,\mathrm{GeV}^2$ we use 27 standard Bernstein averages. To remedy this, we adopt the following approach. For each value of Q^2 we count the number of averages above 9 as duplicate information. The number of duplicates we have altogether is 73 and so for the values of Q^2 for which we have more Bernstein averages than moments, we rescale the error on these averages by $\sqrt{130/(130-73)} = 1.510$. Correspondingly, for the modified Bernstein averages we rescale the errors by a factor of $\sqrt{141/(141-87)} = 1.616$. This rescaling has the effect of suppressing the contribution of the duplicate data points to χ^2, relative to those values of Q^2 for which we have fewer Bernstein averages than moments.

Despite this modification, the use of the naïve form of χ^2 given in Eq. (6.122) rather than the form given in Eq. (6.123) is still incorrect; as a result of this, it will yield unreliable estimates of $\alpha_s(M_Z)$ and $\Lambda_{\overline{\mathrm{MS}}}$. Correspondingly, the errors on these quantities will also be unreliable, and it is possible that they will be underestimated. However, it is worth stressing that the analyses of Refs. [192–194] also used the naïve form of χ^2, and hence the results presented there are equally as unreliable.

6.6.1 Positivity constraints

The fact that xF_3 is a positive definite function, and that the moments are simply integrals over these functions multiplied by a single power of x, means that we can impose certain positivity constraints on the parameters A_n, as follows:

We construct the following matrices from the moments,

$$\hat{\mathcal{M}} = \begin{pmatrix} \mathcal{M}_1 & \mathcal{M}_2 & \cdots & \mathcal{M}_9 \\ \mathcal{M}_2 & \mathcal{M}_3 & & \\ \mathcal{M}_3 & & \ddots & \\ \vdots & & & \\ \mathcal{M}_9 & & & \mathcal{M}_{17} \end{pmatrix}, \tag{6.132}$$

and

$$\Delta\hat{\mathcal{M}} = \begin{pmatrix} \Delta\mathcal{M}_1 & \Delta\mathcal{M}_2 & \cdots & \Delta\mathcal{M}_9 \\ \Delta\mathcal{M}_2 & \Delta\mathcal{M}_3 & & \\ \Delta\mathcal{M}_3 & & \ddots & \\ \vdots & & & \\ \Delta\mathcal{M}_9 & & & \Delta\mathcal{M}_{17} \end{pmatrix}, \tag{6.133}$$

where $\mathcal{M}_n = \mathcal{M}(n;Q^2)$ and $\Delta\mathcal{M}_n = \mathcal{M}_n - \mathcal{M}_{n+1}$.

In order for $\mathcal{M}(n;Q^2)$ to be moments of positive definite functions (as the structure functions must be), the determinants of the above matrices, and of all their minors, must be positive, for all values of Q^2 [193]. Evaluating these determinants at fixed Q^2 will translate to conditions on the parameters A_n. We do not implement these constraints as part of the fitting procedure. Rather, we perform checks on the values of the fitting parameters resulting from the χ^2 minimization in order to ensure that they obey the above constraints.

However, we do impose positivity constraints on the moments themselves. As a result of the determinantal constraints described above, and from the general form of the moments in Eq. (6.1), we can infer that the following inequalities must be satisfied,

$$\mathcal{M}(n;Q^2) > 0, \tag{6.134}$$

$$\mathcal{M}(n;Q^2) \;>\; \mathcal{M}(n+1;Q^2), \tag{6.135}$$

for fixed Q^2. Furthermore, we can implement these constraints by defining our fitting parameters A_n in terms of a new set of parameters and then minimizing χ^2 with respect to these new parameters.

We begin by picking some value of Q_0^2 at which to implement the conditions. We then take the last moment used in the analysis ($n = 20$) and rewrite the constraint in Eq. (6.134) as,

$$\mathcal{M}(20;Q_0^2) \;=\; \left(\hat{A}_{20}\right)^2, \tag{6.136}$$

where $\hat{A}_{20}$ is a real number. The constraints in Eq. (6.135) can also be rewritten as,

$$\mathcal{M}(n;Q_0^2) \;=\; \mathcal{M}(n+1;Q_0^2) + \left(\hat{A}_n\right)^2, \tag{6.137}$$

for $1 \leq n < 20$, where $\hat{A}_n$ are all real numbers. The LHSs of Eqs. (6.136) and (6.137) are simply a fitting parameter times a number. For example, in the case of $n = 2$ and $Q_0^2 = 12.6$ GeV^2 we have,

$$\mathcal{M}(2;12.6\text{ GeV}^2) \;=\; 0.3932\,A_2. \tag{6.138}$$

From this, and equivalent expressions for the rest of the A_n, we can obtain an expression for each A_n in terms of the parameters $\hat{A}_1$ - $\hat{A}_{20}$. This means that we can replace the parameters A_1 - A_{20} with $\hat{A}_1$ - $\hat{A}_{20}$ in the χ^2 function. By doing this and then minimizing with respect to the $\hat{A}_n$ parameters, we can find a minimum in χ^2 for which the constraints in Eqs. (6.134) and (6.135) are automatically satisfied. In effect, the reparameterization embedded in Eqs. (6.136) and (6.137) restricts the parameter space to exclude solutions for which the constraints are not satisfied.

To implement this reparameterization we must choose a value of Q_0^2 at which to impose the constraints, whereas in reality, they must be satisfied for all Q^2. Because of this, we perform the analysis for several different values of Q_0^2 and check that the results remain stable.

6.7 Results

We focus principally on the results from the CORGI analysis in which both odd and even moments are included, and in which we include target mass corrections. This analysis results in a prediction for the QCD scale parameter of,

$$\Lambda^{(5)}_{\overline{\rm MS}} = 219.11^{+23.6}_{-22.1}\ \text{MeV}, \tag{6.139}$$

which corresponds to a value of $\Lambda^{(4)}_{\overline{\rm MS}} = 302.2^{+28.9}_{-27.4}$ MeV in the $N_f = 4$ region. The error on this result can be broken down into four different sources,

$$\Lambda^{(5)}_{\overline{\rm MS}} = 219.11\ ^{+18.36}_{-16.57}\ (\text{stat.})\ \ ^{+8.36}_{-8.17}\ (\text{sys.})\ \ ^{+14.47}_{-13.74}\ (\text{mod.})\ \ \pm 8.97\ (\text{HT})\ \text{MeV}. \tag{6.140}$$

We have used method **I** to obtain experimental values of the averages. The deviation between the results obtained using methods **I** and **IV** is used to evaluate the modelling error since these are the two methods which exhibit the largest deviation.

This result for $\Lambda_{\overline{\rm MS}}$ corresponds to a value of the strong coupling constant (evaluated at the mass of the Z particle) of,

$$\alpha_s(M_Z) = 0.11890^{+0.00191}_{-0.00186}. \tag{6.141}$$

These values show excellent agreement with the current global averages of $\Lambda^{(5)}_{\overline{\rm MS}} = 207.2 \pm 23$ MeV and $\alpha_s(M_Z) = 0.1176 \pm .002$ [32].

The χ^2/d.o.f. for this result is as follows,

$$\begin{aligned} \frac{\chi^2}{\text{d.o.f.}} &= \frac{20.37}{271 - (20+1)} \\ &= 0.0815. \end{aligned} \tag{6.142}$$

In Fig. 6.7 we plot the CORGI predictions for the Bernstein averages (with TMCs included) fitted to the experimental values. Figure 6.8 shows equivalent plots for the modified averages.

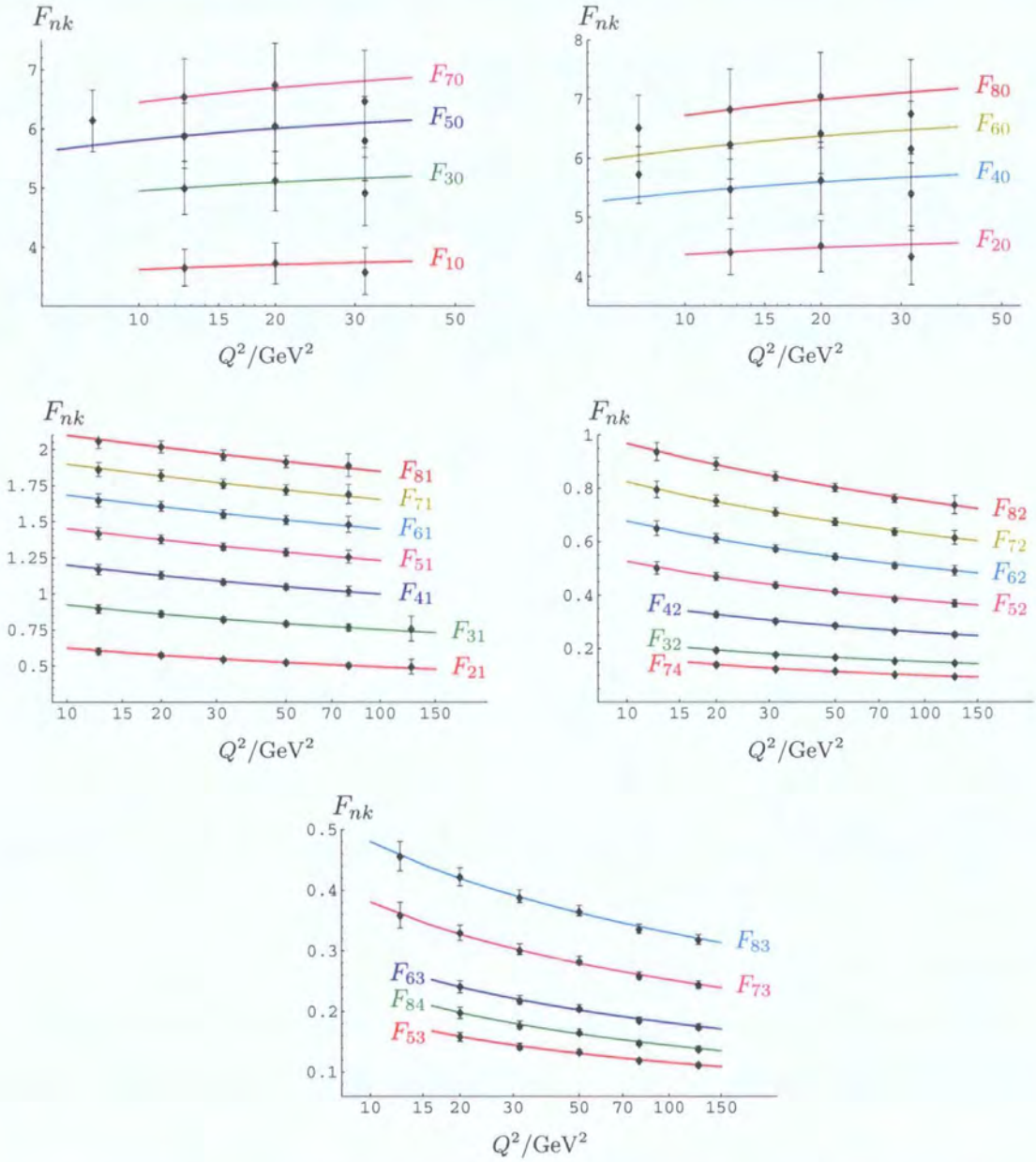


Figure 6.7: CORGI fits for the Bernstein averages of F_3, with TMCs included.

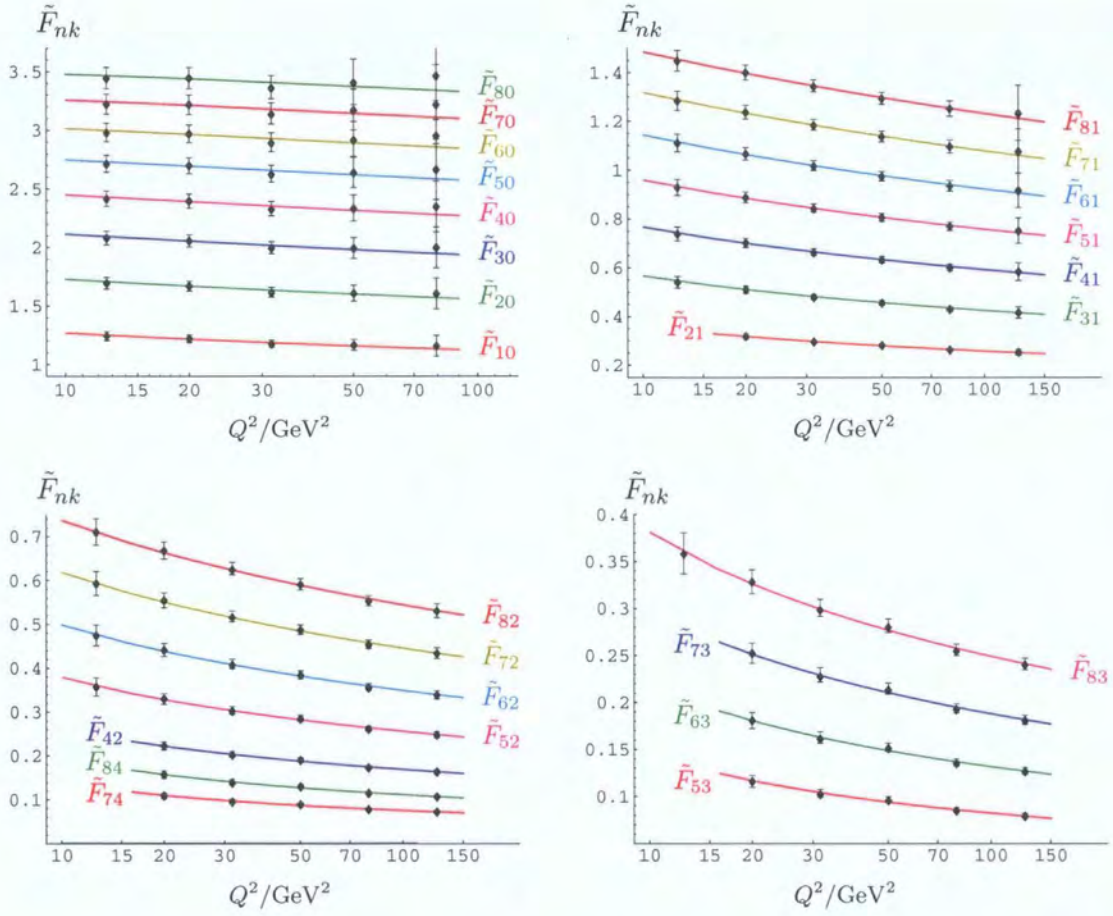


Figure 6.8: CORGI fits for the modified Bernstein averages of F_3, with TMCs included.

	$\Lambda^{(5)}_{\overline{\rm MS}}$(MeV)	$\alpha_s(M_Z)$	χ^2/d.o.f.
All moments	$219.1^{+23.6}_{-22.1}$	$0.1189^{+0.0019}_{-0.0019}$	$20.37/(271-(20+1))$
Odd moments	$210.5^{+35.0}_{-32.7}$	$0.1182^{+0.0029}_{-0.0030}$	$10.94/(130-(10+1))$
Even moments	$229.5^{+64.7}_{-62.1}$	$0.1198^{+0.0048}_{-0.0048}$	$9.24/(141-(10+1))$
All moments: $Q^2 > m_b^2$ only	$232.4^{+34.9}_{-33.2}$	$0.1200^{+0.0027}_{-0.0027}$	$15.59/(228-(20+1))$

Table 6.1: In this table we present the result of the analysis performed using the **CORGI** approach to perturbation theory with target mass corrections included. We compare the results obtained when we include all moments (up to $n = 20$) with those obtained when we restrict the analysis to even or odd moments only. We also show the results from performing the analysis with only data points for which $Q^2 > m_b^2$ ($N_f = 5$) included.

In table 6.1 we present the full set of results from the CORGI analysis. This table shows the results obtained when we include both odd and even moments (standard and modified Bernstein averages) together, and also those obtained when we restrict the analysis to either odd or even moments only. These results allow us to check consistency between the odd, even and 'all moments' analyses.

We can also use the results of the 'odd moments' analysis to check consistency with previous analyses.[2] In the PS analysis of Ref. [193] (in which only the $n = 1, 3, 5, 7, 9, 11$ and 13 moments were included) a value of $\Lambda^{(4)}_{\overline{\rm MS}} = 255 \pm 72$MeV was found, corresponding to $\Lambda^{(5)}_{\overline{\rm MS}} = 178^{+57}_{-55}$MeV, with a value of χ^2/d.o.f. $= 0.007$. Using the same set of moments, the CORGI analysis of Ref. [194] found a value of $\Lambda^{(5)}_{\overline{\rm MS}} = 228^{+35}_{-36}$MeV, although it must be noted that this analysis used incorrect values of the coefficients X_2, and therefore the result must be regarded as unreliable. Both of those results are indeed consistent with the 'odd moments' analysis performed here.

We also perform the analysis in which we restrict the CCFR data to $Q^2 > m_b^2$ only; this will serve as a consistency check on our method of evolving though the b quark threshold. The results of this are also included in table 6.1.

[2] Because the NNLO anomalous dimension coefficients for even n have only recently become available, previous analyses of this kind have included only odd moments. Therefore it is more appropriate to use the results of our 'odd moments' analysis for comparison, rather than the even or 'all moments' results.

n	$A_n^{(4)}$	$\mathcal{M}(n; 8.75\,\mathrm{GeV}^2)$	$\mathcal{M}(n; 12.6\,\mathrm{GeV}^2)$
1	2.346	2.494	2.525
2	0.8814	0.3557	0.3466
3	0.4133	0.1002	9.545×10^{-2}
4	0.2217	3.835×10^{-2}	3.584×10^{-2}
5	0.1292	1.744×10^{-2}	1.603×10^{-2}
6	8.134×10^{-2}	9.048×10^{-3}	8.191×10^{-3}
7	5.241×10^{-2}	4.988×10^{-3}	4.452×10^{-3}
8	3.639×10^{-2}	3.044×10^{-3}	2.681×10^{-3}
9	2.434×10^{-2}	1.826×10^{-3}	1.588×10^{-3}
10	1.822×10^{-2}	1.246×10^{-3}	1.07×10^{-3}
11	1.202×10^{-2}	7.588×10^{-4}	6.438×10^{-4}
12	9.64×10^{-3}	5.677×10^{-4}	4.76×10^{-4}
13	5.935×10^{-3}	3.289×10^{-4}	2.726×10^{-4}
14	5.119×10^{-3}	2.69×10^{-4}	2.204×10^{-4}
15	2.702×10^{-3}	1.355×10^{-4}	1.098×10^{-4}
16	2.535×10^{-3}	1.22×10^{-4}	9.768×10^{-5}
17	9.362×10^{-4}	4.343×10^{-5}	3.44×10^{-5}
18	9.739×10^{-4}	4.376×10^{-5}	3.426×10^{-5}
19	9.807×10^{-10}	4.284×10^{-11}	3.317×10^{-11}
20	7.69×10^{-10}	3.277×10^{-11}	2.509×10^{-11}

Table 6.2: The fitting parameters $A_n^{(4)}$ together with the moments evaluated at $Q^2 = 8.75$ and $12.6\,\mathrm{GeV}^2$.

In table 6.2 we present the values of the parameters A_1 - A_{20} resulting from the CORGI 'all moments' fit with TMCs included, shown in Figs. 6.7 - 6.8. This table also shows the corresponding values of the moments at $Q^2 = 8.75$ and 12.6 GeV^2.

In table 6.3 we compare the CORGI results with those obtained using the PS and EC approaches. We also present results obtained from performing these analyses with and without target mass corrections.

6.8 Summary and conclusions

In this chapter we used three different approaches to perturbation theory to perform a phenomenological analysis of moments of F_3 using the method of Bernstein averages. The three approaches differ in how they deal with the FRS dependence. In the CORGI

		$\Lambda^{(5)}_{\overline{\rm MS}}$(MeV)	$\alpha_s(M_Z)$	χ^2/d.o.f.
CORGI	with TMC	$219.1^{+23.6}_{-22.1}$	$0.1189^{+0.0019}_{-0.0019}$	$20.37/(271-(20+1))$
	no TMC	$280.3^{+24.6}_{-23.4}$	$0.1235^{+0.0017}_{-0.0017}$	$24.76/(272-(18+1))$
PS	with TMC	$200.4^{+25.8}_{-24.8}$	$0.1173^{+0.0023}_{-0.0022}$	$21.73/(273-(20+1))$
	no TMC	$257.5^{+27.8}_{-27.6}$	$0.1219^{+0.0020}_{-0.0020}$	$25.20/(273-(18+1))$
EC	with TMC	$204.5^{+19.9}_{-18.9}$	$0.1177^{+0.0017}_{-0.0017}$	$22.71/(273-(20+1))$
	no TMC	$261.5^{+19.0}_{-18.4}$	$0.1222^{+0.0014}_{-0.0014}$	$26.04/(273-(18+1))$

Table 6.3: In this table we compare the results of the analysis performed with the three different approaches to perturbation theory described in section 6.3, CORGI, PS and EC. We also show the results from these analyses performed with and without target mass corrections.

approach, we allow the FRS invariant quantity $X_0(Q)$ to determine the relationship between M, μ and Q for each moment. In so doing, we automatically resum the subset of terms present in the full perturbative expansion which are RG-predictable at NNLO. In the physical scale approach we set $M = \mu = Q$ and adopt the $\overline{\rm MS}$ scheme for the subtractions in the renormalization *and* factorization procedures. In the effective charge approach, we set $M = \mu$ and act as though we have only one arbitrary scale. Predictions were then obtained by proceeding as we did in the single-scale case in section 1.8.3. We described how predictions are derived in these three approaches and corrected errors in the CORGI results presented in Refs. [39, 194].

We described how target mass and higher twist corrections affect these theoretical predictions and also how we evolve expressions for the moments through the b-quark threshold. We explained how the Bernstein averages method eliminates any potential dependence of the analysis on missing data regions in x and Q^2 and we also described how this method is generalized to treat both odd and even moments. We described the fitting procedure used to extract the optimal values of the QCD scale parameter and how we can implement various constraints which ensure that the results of this fitting are consistent with the structure functions being positive definite functions. We also presented an alternative, and slightly easier method for deriving the FRS invariant quantities X_i.

The results of the CORGI analysis presented in table 6.1 show excellent agreement with the current global average for the strong coupling evaluated at $Q^2 = M_Z^2$ [32]. From this we conclude that CORGI perturbation theory performs well when applied to the analysis of moments.

Although the minimum value of χ^2 per d.o.f. quoted in Eq. (6.142) is an order of magnitude larger than that obtained in Ref. [193], it is still significantly smaller than one would expect. Indeed, in the MRST analysis of Ref. [206], fits are performed to the CCFR $xF_3^{\nu N}$ data, resulting in a value of $\chi^2/\text{d.o.f.} \approx 110/105$. This suggests that the errors on the Bernstein averages in our analysis have been overestimated. However, as discussed in section 6.6, the χ^2 function we are using does not take into account the correlation between data points. Indeed, if correlation was taken into account, one might expect that a more reasonable value of χ^2 would be obtained.

In the χ^2 function we eventually used, the errors on the Bernstein averages were rescaled in order to take into account the 'over-counting of degrees of freedom'. As a result, the number of Bernstein averages is not representative of the true number of degrees of freedom in this particular χ^2 function. Indeed, the Bernstein averages in the plots in Fig. 6.7 can be constructed from just 58 different moments at different values of Q^2 (via Eq. (6.110). Similarly, the modified Bernstein averages in Fig. 6.8 can be built from 53 different moments. Hence,

$$\frac{\chi^2}{\text{d.o.f.}} = \frac{20.37}{111-21} = 0.226, \tag{6.143}$$

is more representative of the true value of $\chi^2/\text{d.o.f.}$ in this approach. This is a more acceptable value (although it is still a factor of 4.6 smaller than the MRST value), however we stress that the *true* minimum in χ^2 can only be determined by taking correlation fully into account. Determining which averages are most strongly correlated, and then excluding these from the fit may be a possible way to do this.

The analyses in which we include only odd or even moments are consistent with each other and with the full (all moments) analysis. Furthermore, in the analysis in which we include all moments, the errors are greatly reduced. This improvement is made possible by the availability of the full NNLO anomalous dimension calculation and represents significant improvement on previous analyses.

Excluding data points for which $Q^2 < m_b^2$ leads to no significant change in the results and from this we conclude that the quark mass threshold method we have applied is suitable to the moment analysis. The error associated with the exclusion of higher twist effects, given in Eq. (6.140), is relatively small, signifying that these effects are not particularly important at scales $Q^2 > 7.6\,\mathrm{GeV}^2$.

We include in the analysis, positivity constraints on the moments (Eqs. (6.134) and (6.135)), via the parameter redefinitions of sections 6.6.1. We find that this implementation has little affect on the prediction of $\Lambda_{\overline{\mathrm{MS}}}$ ($\sim$ 10 Mev), but does make a difference to the predicted values of A_n.

The CORGI predictions for the Bernstein averages (with TMCs included) are plotted in Figs. 6.7 and 6.8 and show excellent agreement with experimental values. This is reflected by the low value of χ^2 associated with this fitting, given in Eq. (6.142).

The results also show consistency between the CORGI, PS and EC approaches. The PS analysis leads to values of $\Lambda_{\overline{\mathrm{MS}}}$ and $\alpha_s(M_Z)$ slightly lower than in the CORGI analysis, and the EC results are lower still. However, this variation is well within the error bars on the associated quantities.

Although the errors quoted in table 6.3 for the CORGI, PS and EC analyses are of comparable magnitude, it is expected that when the theoretical error due to RS dependence and the unknown higher order corrections is included, the errors on the PS quantities will be much larger. By comparison, in the CORGI and EC cases this source of error is greatly reduced.

Inclusion of HT corrections generally results in a small shift in $\Lambda_{\overline{\mathrm{MS}}}$ of about 10 MeV. However, when target mass corrections are included, we see a shift of approximately 60 MeV in the predicted value of $\Lambda_{\overline{\mathrm{MS}}}$, and from this we conclude that these contributions are significant in the case of F_3.

Chapter 7

Analysis of moments of F_2 in electron-proton scattering

In this chapter, we present an analysis of moments of F_2 in electron-proton scattering, within CORGI perturbation theory. We wish to build upon the progress made in the previous chapter by extending the analysis to moments which have a flavour-singlet component. The result of this analysis will be a prediction for the QCD scale parameter, $\Lambda_{\overline{\rm MS}}$ from F_2 data, with the reduced theoretical error associated with the CORGI approach to perturbation theory.

7.1 Introduction

The principal difference between F_2 and F_3 moments is the existence, in the F_2 case, of a singlet contribution to the equations governing their Q^2-dependence. This corresponds to the fact that for the F_2 component of the OPE in Eq. (2.37), there exist lowest-twist operators of each of the types in Eqs. (2.38) - (2.40). The non-singlet (NS) contribution is almost identical to that of F_3 and consequently the Q^2-dependence of this component is described by Eq. (6.65). However, the quark and gluon operators which form the singlet (S) contribution mix under renormalization (factorization). As a result of this, obtaining perturbative predictions for the singlet component of F_2 moments is more troublesome.

In Ref. [192], NNLO expressions for the non-singlet and singlet components of F_2

moments were obtained, and subsequently a physical scale perturbation theory analysis of these moments was carried out. However, we use as *our* starting point the NLO expressions for F_2 derived in Ref. [207], extending them to NNLO. This derivation is more easily adapted to the study of the explicit factorization and renormalization scheme (FRS) dependence of the moments, and therefore allows us to obtain predictions in the CORGI form. In this way, we can perform an analysis of F_2 moments in which the renormalization and factorization scale dependence of the result is reduced, and so improve upon previous analyses.

The existence of a singlet component for F_2 introduces two new sets of free parameters in the expressions for the moments, whose values are to be determined by comparison with data. This effectively triples the number of fitting parameters, and results in the χ^2 minimization procedure becoming more complicated and more subtle. As a consequence of this, the positivity constraints introduced in section 6.6.1, become crucial to the fitting procedure. Only when they are implemented can we eliminate spurious, unphysical minima from χ^2.

As was the case for F_3, the completion of the NNLO calculation of the non-singlet and singlet anomalous dimensions [78,79] extends the set of moments available to us at this order. Hence, we can improve upon previous analyses (limited to even F_2 moments) by including both odd and even moments and extending the highest moment used.

The work presented in this chapter draws mainly on the introductory material covered in chapter 2, and is based on the research carried out in Ref. [208]. The material is organized as follows:

We begin by describing the nature of the singlet contribution to F_2 moments, and detail how we obtain a CORGI prediction for it. We then describe how the method of Bernstein averages may be applied to F_2, including how our approach to modelling F_2 must be modified. We then give details of the fitting procedure and discuss how the positivity constraints, introduced in the last chapter, play a more important role in this analysis. Finally, we present the results of the analysis and discuss what conclusions can be drawn from this work.

7.2 Moments of F_2

In this chapter we are concerned with moments of F_2 in electron-proton scattering, as defined below,

$$\mathcal{M}_2^{ep}(n;Q^2) \;=\; \int_0^1 dx\; x^{n-2} F_2^{ep}(x,Q^2). \tag{7.1}$$

These moments can be decomposed into singlet and non-singlet components,

$$\mathcal{M}_2^{ep}(n;Q^2) \;=\; \mathcal{M}_2^{\rm S}(n;Q^2) + \mathcal{M}_2^{\rm NS}(n;Q^2). \tag{7.2}$$

The origin of the difference between these two components lies in the nature of the operators from which they are derived (Eqs. (2.38) - (2.40)).

The non-singlet component is virtually identical to that of F_3 and therefore we can write down a CORGI prediction for its Q^2-dependence. From Eq. (6.65) we have,

$$\mathcal{M}_2^{\rm NS}(n;Q^2) \;=\; A_n^{\rm NS}\left(\frac{ca_0}{1+ca_0}\right)^{d^{\rm NS}(n)/b}\left(1 + X_2^{\rm NS}(n)a_0^2 + X_3^{\rm NS}(n)a_0^3 + \ldots\right). \tag{7.3}$$

Here, $d^{\rm NS}(n)$ is the first coefficient of the non-singlet anomalous dimension (identical to that used in the F_3 analysis) and $X_2^{\rm NS}(n)$ is an FRS invariant quantity – the F_2 equivalent of Eq. (6.69). Values for $X_2^{\rm NS}(n)$ can be obtained using Eq. (6.70), but with the appropriate F_2 values for the coefficients r_1 and r_2 used (these can be found in Ref. [113]). The coupling a_0 is the same as that defined in Eq. (6.66), with the appropriate scale parameter, $\Lambda_{\mathcal{M}}^{\rm NS}$, given by Eq. (6.68), but with r_1 now being that of the F_2 coefficient function. The coefficients $d^{\rm NS}(n)$, $\Lambda_{\mathcal{M}}^{\rm NS}/\Lambda_{\overline{\rm MS}}$ and $X_2^{\rm NS}(n)$ can be calculated from the results in Refs. [78, 113]; in appendix D we quote their values for $N_f = 4$ and 5, and for the range $2 \leq n \leq 20$. Henceforth, we suppress the notation indicating the n dependence of these parameters.

Knowing that we can obtain an expression for the non-singlet component of the F_2 moments in the CORGI form, we now turn our attention to obtaining similar expressions for the singlet component.

7.2.1 The singlet component

The quark and gluon operators of Eqs. (2.39) and (2.40) mix under renormalization (due to Eq. (2.55)). Consequently, the factorized form of the singlet component (the equivalent of Eq. (6.2)) has contributions from both of these operators,

$$\mathcal{M}_2^{\rm S}(n;Q^2) = \sum_{i=q,g} C^i_{n,\,\rm S}(Q,a(\mu),\mu,M)\langle p|\mathcal{O}^i_{n,\,\rm S}(M)|p\rangle \tag{7.4}$$

$$\equiv \boldsymbol{C}_{n,\,\rm S}(Q,a(\mu),\mu,M)\langle p|\boldsymbol{\mathcal{O}}^{\rm S}_n(M)|p\rangle. \tag{7.5}$$

Here, $\boldsymbol{C}_{n,\,\rm S}(Q,a(\mu),\mu,M)$ and $\langle p|\boldsymbol{\mathcal{O}}^{\rm S}_n(M)|p\rangle$ are the singlet coefficient function and operator matrix element respectively and $|p\rangle$ represents a proton state vector. Boldface type denotes the matrix or vector nature of these components and henceforth, we simplify the notation by defining,

$$\boldsymbol{C}_n \equiv \boldsymbol{C}^{\rm S}_n(Q,a(\mu),\mu,M), \tag{7.6}$$

$$\langle \boldsymbol{\mathcal{O}}_n(M)\rangle \equiv \langle p|\boldsymbol{\mathcal{O}}^{\rm S}_n(M)|p\rangle. \tag{7.7}$$

The perturbative behaviour (specifically the factorization scale dependence) of the operator matrix element is governed by the singlet anomalous dimension equation. From Eqs. (2.55) and (2.59), we have,

$$M\frac{\partial\langle\boldsymbol{\mathcal{O}}_n(M)\rangle}{\partial M} = \boldsymbol{\gamma}^n(a)\langle\boldsymbol{\mathcal{O}}_n(M)\rangle. \tag{7.8}$$

Here, $\boldsymbol{\gamma}^n(a)$ is the singlet anomalous dimension and is a 2×2 matrix, with quark and gluon components,[1]

$$\boldsymbol{\gamma}(a) = \begin{pmatrix} \gamma^{qq}(a) & \gamma^{qg}(a) \\ \gamma^{gq}(a) & \gamma^{gg}(a) \end{pmatrix}. \tag{7.9}$$

Like the non-singlet anomalous dimension, this can be written as an expansion in powers of the coupling, but with each coefficient being a matrix,

$$\boldsymbol{\gamma}(a) = -\boldsymbol{d}a - \boldsymbol{d}_1a^2 - \boldsymbol{d}_2a^3 + \ldots. \tag{7.10}$$

[1]Note that our definition of $\boldsymbol{\gamma}(a)$ is the transpose of that in Ref. [207]. This is due to the fact that in our version of Eq. (7.5), the two components are swapped so that $\boldsymbol{C}_n$ comes first. For other differences between our notation and that adopted in other works, see section D.1.

In the above equation, the relevant coupling is that defined as a function of factorization scale, $a \equiv a(M)$. Similarly, the coefficient function can also be written as an expansion in powers of the coupling,

$$\boldsymbol{C}_n = (C_n^q, C_n^g), \tag{7.11}$$

where,

$$C_n^q = 1 + r_{n,1}^q \tilde{a} + r_{n,2}^q \tilde{a}^2 + \ldots, \tag{7.12}$$

$$C_n^g = r_{n,1}^g \tilde{a} + r_{n,2}^g \tilde{a}^2 + \ldots. \tag{7.13}$$

Here, $\tilde{a} \equiv a(M = \mu)$ is the coupling evaluated at the renormalization scale. Note that the LO coefficient of the gluon component vanishes.

A solution to Eq. (7.8) can be written in the form,

$$\begin{aligned} \langle \mathcal{O}(M) \rangle &= \exp\left\{ \int_0^a \frac{\gamma(x)}{\beta(x)} dx + \int_0^\infty \frac{x\,\boldsymbol{d}}{\beta^{(2)}(x)} dx \right\} \mathbf{A}_n \\ &\equiv \mathbf{W}(a)\mathbf{A}. \end{aligned} \tag{7.14}$$

Here, the components of $\mathbf{A}_n$ are sets of non-perturbative parameters,

$$\mathbf{A}_n = \begin{pmatrix} A_n^q \\ A_n^g \end{pmatrix}; \tag{7.15}$$

they are the analogues of A_n^{NS} and are generated by the factorization process. $\mathbf{W}(a)$ is a 2×2 matrix defined implicitly via Eq. (7.14).

In order to isolate contributions to $\mathbf{W}(a)$ at successive orders of a, we separate it into its LO component and a matrix containing the higher-order components,[2]

$$\mathbf{W}(a) = \exp\left\{ \int_0^a \frac{\gamma(x)}{\beta(x)} dx + \int_0^\infty \frac{x\,\boldsymbol{d}}{\beta^{(2)}(x)} dx \right\} \tag{7.16}$$

$$= \mathbf{V}(a) \exp\left\{ \frac{\boldsymbol{d}}{b} \log\left(\frac{ca}{1+ca} \right) \right\}. \tag{7.17}$$

The form of $\mathbf{V}(a)$ is determined by $\gamma(a)$ and $\beta(a)$; it can be written as an expansion

[2] See Ref. [207] p.252 for the origin of the following derivation.

in powers of a,

$$\mathbf{V}(a) = \mathbf{1} + \mathbf{V}_2 a + \mathbf{V}_3 a^2 + \ldots. \tag{7.18}$$

The coefficients $\mathbf{V}_2$ and $\mathbf{V}_3$ contain the NLO and NNLO information encoded in $\mathbf{W}(a)$, and hence our aim must be to evaluate these matrices. To achieve this, we take the derivative of $\mathbf{W}(a)$ with respect to a using Eqs. (7.16) and (7.17) separately. This yields,

$$\partial_a \mathbf{W}(a) = \frac{\gamma(a)}{\beta(a)} \mathbf{V}(a) \exp\left\{ \frac{\boldsymbol{d}}{b} \log\left(\frac{ca}{1+ca} \right) \right\}, \tag{7.19}$$

$$\partial_a \mathbf{W}(a) = \left(\partial_a \mathbf{V}(a) + \mathbf{V}(a) \frac{\boldsymbol{d}}{ba(1+ca)} \right) \exp\left\{ \frac{\boldsymbol{d}}{b} \log\left(\frac{ca}{1+ca} \right) \right\}. \tag{7.20}$$

Equating these two differentials gives the following equation for $\mathbf{V}(a)$,

$$\partial_a \mathbf{V} + \frac{1}{ba(1+ca)} [\mathbf{V}(a), \boldsymbol{d}] = \left(\frac{\gamma(a)}{\beta(a)} - \frac{\boldsymbol{d}}{ba(1+ca)} \right) \mathbf{V}(a), \tag{7.21}$$

where $[x, y]$ denotes the commutator of x and y. By expanding this expression in powers of a and then equating coefficients, we can obtain equations for $\mathbf{V}_2$ and $\mathbf{V}_3$,

$$b\mathbf{V}_2 + [\mathbf{V}_2, \boldsymbol{d}] = \boldsymbol{d}_1, \tag{7.22}$$

$$2b\mathbf{V}_3 + [\mathbf{V}_3 - c\mathbf{V}_2, \boldsymbol{d}] = \boldsymbol{d}_2 - c\boldsymbol{d}_1 - c_2\boldsymbol{d} + \boldsymbol{d}_1\mathbf{V}_2. \tag{7.23}$$

Solving these equations will yield expressions for the individual components of $\mathbf{V}_2$ and $\mathbf{V}_3$. In order to achieve this it is necessary to diagonalize the first coefficient of the anomalous dimension matrix, $\boldsymbol{d}$. There is some freedom in the choice of diagonalization procedure. We use the matrix,

$$\mathbf{U} = \begin{pmatrix} 1 & -1 \\ \frac{\lambda_- - d^{qq}}{d^{qg}} & \frac{d^{qq} - \lambda_+}{d^{qg}} \end{pmatrix}, \tag{7.24}$$

to diagonalize $\boldsymbol{d}$ thus,

$$\mathbf{U}^{-1} \boldsymbol{d} \mathbf{U} = \begin{pmatrix} \lambda_- & 0 \\ 0 & \lambda_+ \end{pmatrix} \qquad (\lambda_+ > \lambda_- > 0). \tag{7.25}$$

Here, $\lambda_{\mp}$ are the eigenvalues of $\boldsymbol{d}$ and have the form,

$$\lambda_- = \frac{1}{2}\left(d_{qq} + d_{gg} - \sqrt{(d_{gg} - d_{qq})^2 + 4d_{qg}d_{gq}}\right), \tag{7.26}$$

$$\lambda_+ = \frac{1}{2}\left(d_{qq} + d_{gg} + \sqrt{(d_{gg} - d_{qq})^2 + 4d_{qg}d_{gq}}\right). \tag{7.27}$$

We also define $\mathbf{V}(a)$ when acted upon by $\mathbf{U}$ as,

$$\mathbf{U}^{-1}\mathbf{V}(a)\mathbf{U} \equiv \mathbf{V}^D(a), \tag{7.28}$$

and for the coefficients of $\mathbf{V}(a)$,

$$\mathbf{U}^{-1}\mathbf{V}_n\mathbf{U} \equiv \begin{pmatrix} V_n^{--} & V_n^{-+} \\ V_n^{+-} & V_n^{++} \end{pmatrix} \equiv \mathbf{V}_n^D. \tag{7.29}$$

Similarly, for $\boldsymbol{d}_n$ we define,

$$\mathbf{U}^{-1}\boldsymbol{d}_n\mathbf{U} \equiv \begin{pmatrix} d_n^{--} & d_n^{-+} \\ d_n^{+-} & d_n^{++} \end{pmatrix} \equiv \boldsymbol{d}_n^D. \tag{7.30}$$

By acting on equations (7.22) and (7.23) with the diagonalization matrices we obtain four equations for each of the coefficients $\mathbf{V}_2$ and $\mathbf{V}_3$. Solving these yields,

$$\begin{aligned} V_2^{--} &= \frac{d_1^{--}}{b}, & V_2^{-+} &= \frac{d_1^{-+}}{b + \lambda_+ - \lambda_-}, \\ V_2^{+-} &= \frac{d_1^{+-}}{b + \lambda_- - \lambda_+}, & V_2^{++} &= \frac{d_1^{++}}{b}, \end{aligned} \tag{7.31}$$

and,

$$\begin{aligned} V_3^{--} &= \frac{1}{2b}\left(d_2^{--} - cd_1^{--} - c_2\lambda_- + d_1^{--}V_2^{--} + d_1^{-+}V_2^{+-}\right), \\ V_3^{-+} &= \frac{1}{2b + \lambda_+ - \lambda_-}\left(cV_2^{-+}(\lambda_+ - \lambda_-) + d_2^{-+} - cd_1^{-+} + d_1^{--}V_2^{-+} + d_1^{-+}V_2^{++}\right), \\ V_3^{+-} &= \frac{1}{2b + \lambda_- - \lambda_+}\left(cV_2^{+-}(\lambda_- - \lambda_+) + d_2^{+-} - cd_1^{+-} + d_1^{+-}V_2^{--} + d_1^{++}V_2^{+-}\right), \\ V_3^{++} &= \frac{1}{2b}\left(d_2^{++} - cd_1^{++} - c_2\lambda_+ + d_1^{++}V_2^{++} + d_1^{+-}V_2^{-+}\right). \end{aligned} \tag{7.32}$$

These form the vital components of the NLO and NNLO expressions for the singlet moments, as we now show. Inserting factors of $\mathbf{1} = \mathbf{U}\mathbf{U}^{-1}$ into Eq. (7.5), via Eqs. (7.14), (7.17) and (7.29), yields,

$$\mathcal{M}^{\rm S}(n;Q^2) = \boldsymbol{\mathcal{C}}_n \mathbf{U}\,\mathbf{V}^D(a)\,\mathbf{U}^{-1} \exp\left\{\frac{\boldsymbol{d}}{b}\log\left(\frac{ca}{1+ca}\right)\right\}\mathbf{U}\,\mathbf{U}^{-1}\mathbf{A} \tag{7.33}$$

$$= \bar{\boldsymbol{\mathcal{C}}}_n\,\mathbf{V}^D(a) \begin{pmatrix} \left(\frac{ca}{1+ca}\right)^{\lambda_-/b} & 0 \\ 0 & \left(\frac{ca}{1+ca}\right)^{\lambda_+/b} \end{pmatrix} \bar{\mathbf{A}}. \tag{7.34}$$

Here,

$$\bar{\mathbf{A}} \equiv \mathbf{U}^{-1}\mathbf{A} \equiv \begin{pmatrix} A^- \\ A^+ \end{pmatrix}, \tag{7.35}$$

and $\bar{\boldsymbol{\mathcal{C}}}_n$ is a transformation of the coefficient function,

$$\bar{\boldsymbol{\mathcal{C}}} = \boldsymbol{\mathcal{C}}\,\mathbf{U} \tag{7.36}$$

$$= \left(1 + r_1^-\,\tilde{a} + r_2^-\,\tilde{a}^2 + \ldots, -1 - r_1^+\,\tilde{a} - r_2^+\,\tilde{a}^2 - \ldots\right), \tag{7.37}$$

with,

$$r_i^{\mp} = r_i^q + r_i^g \frac{\lambda_{\mp} - d^{qq}}{d^{qg}}. \tag{7.38}$$

At this point we set $a = \tilde{a}$ and expand equation (7.34) in powers of a. Although it is necessary to make this identification at this stage, it will still be possible to recover expressions in the CORGI form from the eventual result, through the method outlined in section 6.4.

The diagonalization carried out in Eq. (7.34) causes the singlet moments to decouple into (+) and (−) components. Expanding Eq. (7.34) gives,

$$\mathcal{M}^{\rm S}(n;Q^2) = \bar{A}^- \left(\frac{ca}{1+ca}\right)^{\lambda_-/b} \left(1 + R_1^- a + R_2^- a^2\right)$$

$$-\bar{A}^+ \left(\frac{ca}{1+ca}\right)^{\lambda_+/b} \left(1 + R_1^+ a + R_2^+ a^2\right) \tag{7.39}$$

$$= \mathcal{M}^-(n;Q^2) - \mathcal{M}^+(n;Q^2), \tag{7.40}$$

where,

$$R_1^{\mp} = r_1^{\mp} + V_2^{\mp\mp} - V_2^{\pm\mp}, \tag{7.41}$$

$$R_2^{\mp} = r_2^{\mp} + r_1^{\mp} V_2^{\mp\mp} - r_1^{\pm} V_2^{\pm\mp} + V_3^{\mp\mp} - V_3^{\pm\mp}. \tag{7.42}$$

The constants $A_n^{\mp}$ are related to the original parameters by an FRS independent transformation and hence, are equally valid fitting parameters of the analysis. Setting $Q = M$ at this stage, and then adopting the three loop coupling of Eq. (1.118), would correspond to the PS approach, detailed in sections 1.8.2 and 6.3.2. However, we wish to adapt Eq. (7.39) to CORGI perturbation theory, and we describe how this is done in the next section.

7.3 CORGI predictions for F_2 moments

Equation (7.39) can be rewritten in the form of two effective charges $\tilde{R}^+(a)$ and $\tilde{R}^-(a)$ [41],

$$\mathcal{M}^{\rm S}(n;Q^2) = \bar{A}^- \left(c\tilde{R}^-(a)\right)^{\lambda_-/b} - \bar{A}^+ \left(c\tilde{R}^+(a)\right)^{\lambda_+/b}, \tag{7.43}$$

where,

$$\tilde{R}^{\pm}(a) = a + \tilde{R}_1^{\pm} a^2 + \tilde{R}_2^{\pm} a^3 + \ldots. \tag{7.44}$$

However, we can relate this form of the moments to the CORGI approach by instead rewriting this in terms of two new effective charges, $\hat{R}^{\mp}(a)$,

$$\mathcal{M}^{\rm S}(n;Q^2) = A_n^- \left(\frac{c\hat{R}^-(a)}{1+c\hat{R}^-(a)}\right)^{\lambda_-/b} - A_n^+ \left(\frac{c\hat{R}^+(a)}{1+c\hat{R}^+(a)}\right)^{\lambda_+/b} \tag{7.45}$$

with

$$\hat{R}^{\pm}(a) = a + \hat{R}_1^{\pm} a^2 + \hat{R}_2^{\pm} a^3 + \ldots. \tag{7.46}$$

The coefficients $\hat{R}_i^{\mp}$ and $\tilde{R}_i^{\mp}$ can be obtained by comparing Eqs. (7.43) and (7.45) with Eq. (7.39). Using the method described in section 6.4, we can then write,

$$\mathcal{M}^{\rm S}(n;Q^2) = A_n^- \left(\frac{ca_0^-}{1+ca_0^-}\right)^{\lambda_-/b} \left(1 + X_2^- \left(a_0^-\right)^2 + \ldots\right)$$

$$- A_n^+ \left(\frac{ca_0^+}{1+ca_0^+}\right)^{\lambda_+/b} \left(1 + X_2^+ (a_0^+)^2 + \ldots\right). \qquad (7.47)$$

The coefficients $X_2^{\mp}$ are obtainable from Eq. (6.97) and they have the form,

$$\begin{aligned} X_2^{\mp} &= \frac{\lambda_{\mp}}{b}\left(\hat{R}_2^{\mp} - \left(\hat{R}_1^{\mp}\right)^2 - c\hat{R}_1^{\mp} + c_2\right) \\ &= R_2^{\mp} + c_2\frac{\lambda_{\mp}}{b} - \frac{\left(R_1^{\mp}\right)^2}{2}\left(1 + \frac{b}{\lambda_{\mp}}\right). \end{aligned} \qquad (7.48)$$

The coupling in Eq. (7.47) is the familiar 't Hooft coupling of Eq. (6.66). However the scale $\Lambda_{\mathcal{M}}$ has a different form for the (+) and (−) component. The relation between $\Lambda_{\mathcal{M}}^{\mp}$ and $\Lambda_{\overline{\rm MS}}$ can be obtained from Eq. (6.101) and is given in the summary below.

From the above results we can write down a CORGI prediction for both the non-singlet and singlet components of the F_2 moments. The full result has the form,

$$\begin{aligned} \mathcal{M}(n;Q^2) &= A_n^{\rm NS}\left(\frac{ca_0^{\rm NS}}{1+ca_0^{\rm NS}}\right)^{d^{\rm NS}/b}\left(1 + X_2^{\rm NS}(a_0^{\rm NS})^2 + \ldots\right) \\ &+ A_n^-\left(\frac{ca_0^-}{1+ca_0^-}\right)^{\lambda_-/b}\left(1 + X_2^-(a_0^-)^2 + \ldots\right) \\ &- A_n^+\left(\frac{ca_0^+}{1+ca_0^+}\right)^{\lambda_+/b}\left(1 + X_2^+(a_0^+)^2 + \ldots\right). \end{aligned} \qquad (7.49)$$

Due to the relation between the scales $\Lambda_{\mathcal{M}}$ and $\Lambda_{\overline{\rm MS}}$ being dependent on which component of Eq. (7.49) we are referring to, the coupling has a different form in each case. Thus,

$$a_0^i(Q) = \frac{-1}{c\left[1 + W_{-1}(z^i(Q))\right]}, \qquad z^i(Q) = -\frac{1}{\rm e}\left(\frac{Q}{\Lambda_{\mathcal{M}}^i}\right)^{-b/c}, \qquad (7.50)$$

with i =NS, (−) and (+). Finally, the scales $\Lambda_{\mathcal{M}}^i$ are given by,

$$\Lambda_{\mathcal{M}}^{\rm NS} = \tilde{\Lambda}_{\overline{\rm MS}} \exp\left\{\frac{d_1^{\rm NS}}{d^{\rm NS}b} + \frac{r_1^{\rm NS}}{d^{\rm NS}}\right\}, \qquad (7.51)$$

$$\Lambda_{\mathcal{M}}^{\mp} = \tilde{\Lambda}_{\overline{\rm MS}} \exp\left\{\frac{R_1^{\mp}}{\lambda_{\mp}}\right\}. \qquad (7.52)$$

The coefficients in Eq. (7.49) and those in Eq. (7.39) are derived from the coefficients

of the singlet and non-singlet anomalous dimensions [78, 79] and the F_2 coefficient functions [113]. We quote their values in appendix D, for $N_f = 4$ and $N_f = 5$, and for the range $2 \leq n \leq 20$.

7.4 Additional effects

As was the case in the F_3 analysis, it is necessary to include a number of other effects to supplement the perturbative prediction of Eq. (7.49). These are effects due to the b-quark mass threshold, target mass corrections and higher twist operators.

Matching of the coupling at the mass of the b-quark leads to a relation between $\Lambda^{(4)}_{\overline{\rm MS}}$ and $\Lambda^{(5)}_{\overline{\rm MS}}$, as given by Eq. (5.24). From this, we can derive values for the scales $\Lambda^i_{\mathcal{M}}$ in the two regions. Continuity of the moments themselves is ensured by relating the A^i_n parameters in these two regions by,

$$A_n^{i,(5)} = \left(\frac{A_n^i}{\mathcal{M}^i(n;m_b^2)}\bigg|_{N_f=5} \right) \left(\frac{\mathcal{M}^i(n;m_b^2)}{A_n^i}\bigg|_{N_f=4} \right) A_n^{i,(4)}, \tag{7.53}$$

where i =NS, (+) and (−).

Target mass corrections for F_2 have the form [198, 199],

$$\mathcal{M}^{\rm TMC}(n;Q^2) = \mathcal{M}(n;Q^2) + \frac{n(n-1)}{n+2}\frac{m_p^2}{Q^2}\mathcal{M}(n+2;Q^2), \tag{7.54}$$

where m_p is the mass of the proton. We apply this expression to both the non-singlet *and* singlet components.

Higher twist effects are estimated by means of the following expression,

$$\mathcal{M}^{\rm HT}(n;Q^2) = n\left(A^{\rm HT}\frac{\Lambda^2_{\overline{\rm MS}}}{Q^2}\right)\mathcal{M}(n;Q^2). \tag{7.55}$$

We only include this term in order to estimate the error associated with our poor understanding of these effects. The deviation between the result of the analysis with and without Eq. (7.55) included will form the 'HT error'. The final quoted results will be obtained with this term omitted.

7.5 Bernstein averages of F_2

The definition of the Bernstein averages of F_2 is identical to that for F_3,

$$F_{nk}(Q^2) = \int_0^1 p_{nk}(x^2)F_2(x,Q^2)dx, \tag{7.56}$$

$$\tilde{F}_{nk}(Q^2) = \int_0^1 \tilde{p}_{nk}(x^2)F_2(x,Q^2)dx. \tag{7.57}$$

We wish to study both odd and even moments and so we include both standard and modified averages in our analysis. The functions $p_{nk}(x)$ and $\tilde{p}_{nk}(x)$ are defined in Eqs. (6.103) and (6.112) and the ranges of x enhanced by the averages are the same as before (Eqs. (6.109) and (6.116)). We also use the same acceptance criteria as that described in section 6.5.

7.5.1 Modelling F_2

To obtain values of the averages from experimental data it is necessary to model the structure functions over the whole range of x, for each value of Q^2. The result is expected to be independent of this modelling, due to the nature of the Bernstein polynomials. Nonetheless, we must first choose a method of extrapolating and interpolating the moments. We use the methods **I** - **IV** described in section 6.5, modified in the following ways:

I′ To account for the fact that F_2 has a singlet contribution, we modify the fitting function of Eq. (6.111) to,

$$F_2(x) = F_{\rm S}(x,Q^2) + F_{\rm NS}(x,Q), \tag{7.58}$$

where,

$$F_{\rm S}(x,Q^2) = (A\,x^{-.44} + B)(1-x)^{\nu}, \tag{7.59}$$

$$F_{\rm NS}(x,Q^2) = C\,x^{1/2}(1-x)^{\mu}. \tag{7.60}$$

The parameters $\{A, B, C, \mu, \nu\}$ are obtained by fitting to the data. A theoretical justification for the use of this function can be found in Ref. [204]. However, we once again stress that the final result will be independent of the fitting method

used. We find that Eqs. (7.58) - (7.60) fit the data well in almost all cases.

II′ The only modification we make to method **II** is to change the extrapolation between $x = 0$ and the first data point. Instead of extrapolating to the point $F_2(x = 0) = 0$, we use the gradient between the two lowest-x data points to extrapolate back to $x = 0$. This makes method **II′** consistent with **I′**.

III′ For method **III** the only modification we make is to use the fitting function of Eqs. (7.58) - (7.60) instead of Eq. (6.111).

IV′ Method **IV** is left unaltered.

In method **I′**, the experimental averages can be obtained directly, once we have fitted values of A, B, C, μ and ν at each value of Q^2. Substituting Eq. (7.58) into Eqs. (7.56) and (7.57), and using the definition of the beta function given in Eq. (B.5) yields,

$$F_{nk}^{(2)} = \frac{2\Gamma(n+\frac{3}{2})}{\Gamma(k+\frac{1}{2})}\sum_{l=0}^{n-k}\frac{(-1)^l}{l!(n-k-l)!}\Big[A\ \mathrm{Beta}(2(k+l)+0.56,\nu+1) + B\ \mathrm{Beta}(2(k+l)+1,\nu+1) + C\ \mathrm{Beta}(2(k+l)+1.5,\mu+1)\Big], \quad (7.61)$$

$$\tilde{F}_{nk}^{(2)} = \frac{2\Gamma(n+2)}{\Gamma(k+1)}\sum_{l=0}^{n-k}\frac{(-1)^l}{l!(n-k-l)!}\Big[A\ \mathrm{Beta}\,(2(k+l)+1.56,\nu+1) + B\ \mathrm{Beta}(2(k+l)+2,\nu+1) + C\ \mathrm{Beta}(2(k+l)+2.5,\mu+1)\Big]. \quad (7.62)$$

For method **III′**, the averages are obtained by evaluating the integrals in Eqs. (7.56) and (7.57) between the appropriate limits of integration, and in methods **II′** and **IV′**, the averages are obtained in the same way as they were in the F_3 analysis.

7.5.2 Experimental input

Data for F_2^{ep} is available from a number of different sources. We use data from SLAC [209], BCDMS [210], E665 [211], ZEUS [212], H1 [213, 214] and NMC [215]. The principal difficulty in organizing this data into a practicable form stems from the fact that different experiments have chosen different sets of Q^2 values at which to perform

measurements. If two experiments have measured F_2 at values of Q^2 that are in the immediate vicinity of each other, then we amalgamate these data sets and assign them to a single Q^2 value. Thus, in the final result, we include an error associated with the uncertainty in the value of Q^2. In this way we can extend the range of x covered at each value of Q^2.

7.6 The fitting procedure and positivity constraints

We use χ^2 minimization to fit the predictions of Eq. (7.49) to experimental data. The χ^2 function has the following form,

$$\chi^2\left(\{A_n^{\rm NS}\},\{A_n^-\},\{A_n^+\},\Lambda_{\overline{\rm MS}},A_{\rm HT}\right) = \sum_{\substack{n=1\\n>k}}^{\frac{1}{2}(n_{\rm max}-1)}\sum_{k=0}\sum_{Q^2\in\{Q_i^2\}}\left(\frac{F_{nk}^{\rm theory}(Q^2)-F_{nk}^{\rm exp}(Q^2)}{\sigma F_{nk}^{\rm exp}(Q^2)}\right)^2 + \sum_{\substack{n=1\\n>k}}^{\frac{1}{2}(n_{\rm max}-1)}\sum_{k=0}\sum_{Q^2\in\{Q_i^2\}}\left(\frac{\tilde{F}_{nk}^{\rm theory}(Q^2)-\tilde{F}_{nk}^{\rm exp}(Q^2)}{\sigma \tilde{F}_{nk}^{\rm exp}(Q^2)}\right)^2, \tag{7.63}$$

where $\{Q_i^2\}$ is the set of values of Q^2 for which we have experimental values for the averages. The errors on the averages, $\sigma F_{nk}^{\rm exp}(Q^2)$, are obtained by summing statistical, systematic and modelling errors in quadrature.

As was the case in the F_3 analysis, the χ^2 function given in Eq. (7.63) does not take into account the correlation between data for the averages. Hence it is considered rather too naïve and will lead to an unreliable extraction of $\alpha_s(M_Z)$ and $\Lambda_{\overline{\rm MS}}$. However, the covariance matrices for the F_2 averages are, like their F_3 counterparts, ill-conditioned. As a result, it is impossible to perform a χ^2 minimization with the correlation taken into account.

In light of this we settle for the approach taken in chapter 6 of rescaling the errors on the averages in order to take into account the 'over-counting of degrees of freedom'. This results in the experimental errors on the Bernstein averages and modified Bernstein averages being rescaled by factors of $\sqrt{789/(789-524)} = 1.73$ and $\sqrt{1058/(1058-694)} = 1.70$ respectively.

The lowest moment we include is $n = 2$ and the highest is $n_{\max} \equiv 17$. Including moments higher than this leads to no significant increase in the number of acceptable averages, and hence will not lead to any increase on the precision of the prediction for $\Lambda_{\overline{\rm MS}}$. Limiting the number of moments also ensures that the computing time taken to minimize χ^2 is not unreasonably large. When target mass corrections are included, we have a total of $3n_{\max} + 3 + 1 = 55$ fitting parameters, consisting of 18 parameters for each of the components of Eq. (7.49) plus the scale parameter $\Lambda_{\overline{\rm MS}}$. Including higher twist corrections expands this set to 56 parameters.

The large number of free parameters means that the minimization of χ^2 is a rather complicated exercise. The *most* complicated part of χ^2 is its dependence on $\Lambda_{\overline{\rm MS}}$, and things are simplified greatly if we fix this parameter and minimize with respect to the remaining parameters. In this case, χ^2 is simply an order-2 multinomial in A_n^i. Consequently, minimization can be achieved by setting to zero the differentials of χ^2 with respect to each of the fitting parameters, and then solving the resultant system of simultaneous equations. This is still a far from trivial exercise. However it is an order of magnitude simpler than minimization with respect to $\Lambda_{\overline{\rm MS}}$. In light of this, we adopt the following approach to minimizing χ^2:

We minimize χ^2 for fixed $\Lambda_{\overline{\rm MS}}$ via the method described above, and then repeat this procedure for a range of $\Lambda_{\overline{\rm MS}}$ values. The result of this is a set of values of the minimum possible χ^2, for each value of $\Lambda_{\overline{\rm MS}}$. From this we can directly identify the optimal value of $\Lambda_{\overline{\rm MS}}$ (the minimum in χ^2), as well as the level of precision associated with this value (the values of $\Lambda_{\overline{\rm MS}}$ corresponding to $\chi^2_{\min} + 1$).

Positivity constraints

The increase in the number of fitting parameters (relative to the F_3 analysis) increases the importance of implementing the positivity constraints on the moments. Without them, it is possible that spurious, unphysical minima may occur in the χ^2 function, and we wish to eliminate these.

In the F_3 analysis we performed a redefinition of our fitting parameters, which had the effect of restricting the parameter space such that certain positivity constraints on the moments (Eqs. (6.134) and (6.135)) were satisfied automatically. We then checked that

the values of A_n resulting from the fit, satisfied the full set of determinantal constraints (see Eqs. (6.132) and (6.133)). In the F_2 analysis, we shall proceed as we did for F_3, but we wish to make this more sophisticated by including some of the determinantal constraints as part of the fitting procedure.

7.6.1 Positivity of the singlet and non-singlet moments

The moments of F_2 have two separate components, each of which must independently satisfy constraints of the type described in section 6.6.1. Therefore we have,

$$\mathcal{M}^{\rm NS}(n;Q^2) > 0, \tag{7.64}$$

$$\mathcal{M}^{\rm S}(n;Q^2) > 0, \tag{7.65}$$

$$\mathcal{M}^{\rm NS}(n;Q^2) > \mathcal{M}^{\rm NS}(n+1;Q^2), \tag{7.66}$$

$$\mathcal{M}^{\rm S}(n;Q^2) > \mathcal{M}^{\rm S}(n+1;Q^2). \tag{7.67}$$

We discuss how *these* constraints are implemented first, before dealing with the rather more complicated determinantal constraints.

The non-singlet component

For the non-singlet component we proceed as we did for F_3 by defining a new set of parameters, $\hat{A}_n^{\rm NS}$,

$$\mathcal{M}^{\rm NS}(n,Q_0^2) \equiv \sum_{i=n}^{n_{\rm max}+2} \left(\hat{A}_n^{\rm NS}\right)^2, \tag{7.68}$$

where Q_0^2 is some reference energy scale. This redefinition ensures that Eqs. (7.64) and (7.66) are satisfied automatically for all real values of $\hat{A}_n^{\rm NS}$ at $Q^2 = Q_0^2$. In analogy with Eq. (6.138), for fixed n the LHS of the above equation is simply a number times $A_n^{\rm NS}$. Hence we can replace the set of parameters A_2 - $A_{n_{\rm max}+2}$ with $\hat{A}_2$ - $\hat{A}_{n_{\rm max}+2}$ and minimize χ^2 with respect to the latter set. In this way, we can guarantee that the results of the χ^2 minimization will correspond to values of the moments which satisfy Eqs. (7.64) and (7.66) at $Q^2 = Q_0^2$.

The singlet component

The singlet component can be split into (+) and (−) components. From Eq. (7.47),

$$\mathcal{M}^{\rm S}(n,Q^2) \;=\; \mathcal{M}^{-}(n,Q^2) - \mathcal{M}^{+}(n,Q^2). \tag{7.69}$$

We can use the small and large Q^2 behaviour of the moments to infer constraints on the individual (+) and (−) components. At large Q^2, the smallest power of a in Eq. (7.49) will dominate the behaviour of the moments. From appendix D we can see that $\lambda_- \leq d^{\rm NS} \leq \lambda_+$, for all n and therefore the dominant term in Eq. (7.49) is the LO (−) component,

$$\mathcal{M}(n;Q^2) \underset{Q^2\to\infty}{\approx} \mathcal{M}^{-}(n;Q^2). \tag{7.70}$$

Hence, in order for the moments to remain positive in the limit of asymptotic freedom we must demand,

$$\mathcal{M}^{-}(n;Q^2) \;>\; 0, \tag{7.71}$$

$$\Longrightarrow\; A_n^{-} \;>\; 0, \tag{7.72}$$

for all Q^2.

At small Q^2, the highest power of a will dominate, and hence,

$$\mathcal{M}(n;Q^2) \underset{Q^2\to 0}{\approx} -A_n^{+} c^{\lambda_+/b} X_2^{+} a_0^{2+\lambda_+/b}. \tag{7.73}$$

This would appear to suggest that we should demand that $A_n^{+} < 0$. However, due to the uncertainty in the sign of the beyond NNLO coefficients, we cannot draw any firm conclusions about the sign of A_n^{+}.

In light of the above constraint on A_n^{-}, we now define two new sets of parameters, $\hat{A}_n^{\rm S}$ and $\hat{A}_n^{-}$,

$$\mathcal{M}^{-}(n;Q_0^2) \;=\; \left(\hat{A}_n^{-}\right)^2, \tag{7.74}$$

$$\mathcal{M}^{+}(n;Q_0^2) \;=\; \mathcal{M}^{-}(n;Q_0^2) - \sum_{i=n}^{n_{\rm max}+2} \left(\hat{A}_i^{\rm S}\right)^2. \tag{7.75}$$

The above equations ensure that Eqs. (7.65), (7.67) and (7.71) are satisfied for all real values of the new parameters, at Q_0^2. Using the relationship between $\mathcal{M}^{\pm}(n;Q^2)$ and $A_n^{\pm}$, together with the above equations, we can replace the original fitting parameters $A_2^{\pm}$ - $A_{n_{\max}+2}^{\pm}$ with $\hat{A}_2^{\rm S}$ - $\hat{A}_{n_{\max}+2}^{\rm S}$ and $\hat{A}_2^{-}$ - $\hat{A}_{n_{\max}+2}^{-}$ in χ^2, for all n. Minimizing with respect to these new parameters will force the minimum in χ^2 to correspond to values of the singlet component of the moments which satisfy Eqs. (7.65), (7.67) and (7.71).

7.6.2 Determinantal constraints

In addition to the positivity constraints on the values of the moments, demanding that the structure functions are positive definite functions implies a further set of constraints, in the form of inequality relations on the determinants of matrices of the type in Eqs. (6.132) and (6.133). In the case of F_2, we can apply these constraints to the singlet and non-singlet components separately. Furthermore, we find that the crucial components of these constraints involve the following matrices,

$$\begin{pmatrix} \mathcal{M}^i(n;Q^2) & \mathcal{M}^i(n+1;Q^2) \\ \mathcal{M}^i(n+1;Q^2) & \mathcal{M}^i(n+2;Q^2) \end{pmatrix}, \tag{7.76}$$

where i =NS or S. We find that demanding that these matrices have positive determinants, causes most of the rest of the constraints to be satisfied automatically. In order to implement the above constraints as part of the fitting procedure, we adopt a parameter redefinition similar to that of the last section. In practice, we only implement Eq. (7.76) for the first two values of n. Therefore we have,

$$\det \begin{vmatrix} \mathcal{M}^i(2;Q^2) & \mathcal{M}^i(3;Q^2) \\ \mathcal{M}^i(3;Q^2) & \mathcal{M}^i(4;Q^2) \end{vmatrix} > 0, \qquad \det \begin{vmatrix} \mathcal{M}^i(3;Q^2) & \mathcal{M}^i(4;Q^2) \\ \mathcal{M}^i(4;Q^2) & \mathcal{M}^i(5;Q^2) \end{vmatrix} > 0, \tag{7.77}$$

where i = S or NS.

The non-singlet component

For the non-singlet component, we define the non-singlet version of the determinants in Eq. (7.77) in terms of two new parameters, $\tilde{A}_2^{\rm NS}$ and $\tilde{A}_3^{\rm NS}$. We then apply the necessary

constraints at the reference energy Q_0^2,

$$\left(\tilde{A}_2^{\rm NS}\right)^2 \equiv \mathcal{M}^{\rm NS}(2;Q_0^2)\mathcal{M}^{\rm NS}(4;Q_0^2) - \left(\mathcal{M}^{\rm NS}(3;Q_0^2)\right)^2 \geq 0, \tag{7.78}$$

$$\left(\tilde{A}_3^{\rm NS}\right)^2 \equiv \mathcal{M}^{\rm NS}(3;Q_0^2)\mathcal{M}^{\rm NS}(5;Q_0^2) - \left(\mathcal{M}^{\rm NS}(4;Q_0^2)\right)^2 \geq 0. \tag{7.79}$$

Rearranging the above yields,

$$\mathcal{M}^{\rm NS}(2;Q_0^2) = \frac{\left(\mathcal{M}^{\rm NS}(3;Q_0^2)\right)^2 + \left(\tilde{A}_2^{\rm NS}\right)^2}{\mathcal{M}^{\rm NS}(4;Q_0^2)}, \tag{7.80}$$

$$\mathcal{M}^{\rm NS}(3;Q_0^2) = \frac{\left(\mathcal{M}^{\rm NS}(4;Q_0^2)\right)^2 + \left(\tilde{A}_3^{\rm NS}\right)^2}{\mathcal{M}^{\rm NS}(5;Q_0^2)}. \tag{7.81}$$

The LHSs of Eqs. (7.80) and (7.81) are simply multiples of the parameters $A_2^{\rm NS}$ and $A_3^{\rm NS}$ respectively. Hence we can use these equations together with Eq. (7.68) to replace the set of parameters $A_2^{\rm NS}$ - $A_{n_{\rm max}+2}^{\rm NS}$ with a new set: $\tilde{A}_2^{\rm NS}$, $\tilde{A}_3^{\rm NS}$ and $\hat{A}_4^{\rm NS}$ - $\hat{A}_{n_{\rm max}+2}^{\rm NS}$. Minimizing with respect to this new set of parameters will restrict the parameter space such that Eqs. (7.64), (7.66) and (7.77) are satisfied automatically at $Q^2 = Q_0^2$.

The singlet component

For the singlet component we proceed in the same manner by defining the singlet determinants of Eq. (7.77), evaluated at Q_0^2, in terms of two new parameters, $\tilde{A}_2^{\rm s}$ and $\tilde{A}_3^{\rm s}$,

$$\left(\tilde{A}_2^{\rm s}\right)^2 \equiv \mathcal{M}^{\rm s}(2;Q_0^2)\mathcal{M}^{\rm s}(4;Q_0^2) - \left(\mathcal{M}^{\rm s}(3;Q_0^2)\right)^2 \geq 0, \tag{7.82}$$

$$\left(\tilde{A}_3^{\rm s}\right)^2 \equiv \mathcal{M}^{\rm s}(3;Q_0^2)\mathcal{M}^{\rm s}(5;Q_0^2) - \left(\mathcal{M}^{\rm s}(4;Q_0^2)\right)^2 \geq 0. \tag{7.83}$$

Rearranging the above yields,

$$\mathcal{M}^{\rm s}(2;Q_0^2) = \frac{\left(\mathcal{M}^{\rm s}(3;Q_0^2)\right)^2 + \left(\tilde{A}_2^{\rm s}\right)^2}{\mathcal{M}^{\rm s}(4;Q_0^2)}, \tag{7.84}$$

$$\mathcal{M}^{\rm s}(3;Q_0^2) = \frac{\left(\mathcal{M}^{\rm s}(4;Q_0^2)\right)^2 + \left(\tilde{A}_3^{\rm s}\right)^2}{\mathcal{M}^{\rm s}(5;Q_0^2)}. \tag{7.85}$$

The LHSs of these equations are simply linear combinations of $A_2^{\pm}$ and $A_3^{\pm}$ respectively. Hence we can use these equations together with Eqs. (7.74) and (7.75) to replace the parameters $A_2^{\pm}$ - $A_{n_{\max}+2}^{\pm}$ with $\tilde{A}_2^{\rm s}$, $\tilde{A}_3^{\rm s}$, $\hat{A}_4^{\rm s}$ - $\hat{A}_{n_{\max}+2}^{\rm s}$ and $\hat{A}_2^{-}$ - $\hat{A}_{n_{\max}+2}^{-}$. Minimizing with respect to this new set of parameters will restrict the parameter space such that Eqs. (7.65), (7.67) and the singlet component of Eq. (7.77) are satisfied automatically at $Q^2 = Q_0^2$.

In summary, by defining new sets of fitting parameters, we can restrict the parameter space such that chosen sets of positivity conditions are satisfied automatically at a given value of Q_0^2.

7.7 Results

In Fig. 7.1 we plot the CORGI predictions for the Bernstein averages (with TMCs included) fitted to experimental values. Figure 7.2 shows equivalent plots for the modified averages. The plots show excellent agreement between theory and experiment, and result in a prediction for the QCD scale parameter of,

$$\Lambda_{\overline{\rm MS}}^{(5)} = 267.3^{+23.4}_{-19.1}\ {\rm MeV}. \tag{7.86}$$

The corresponding value of the strong coupling constant (evaluated at the mass of the Z particle) is,

$$\alpha_s(M_Z) = 0.1226^{+0.0012}_{-0.0017}. \tag{7.87}$$

These values are within 2 – 3 standard deviations of the current global averages of $\Lambda_{\overline{\rm MS}}^{(5)} = 207.2 \pm 23$ MeV and $\alpha_s(M_Z) = 0.1176 \pm 0.002$ [32]. However, this deviation is not unreasonable, due to the differences between the analysis performed here and those which contributed to the results in Ref. [32].

The corresponding value of χ^2 per degree of freedom is,

$$\begin{aligned} \frac{\chi^2}{\text{d.o.f.}} &= \frac{752.665}{1847 - (54+1)} \\ &= 0.420, \end{aligned} \tag{7.88}$$

	$\Lambda^{(5)}_{\overline{\rm MS}}$(MeV)	$\alpha_s(M_Z)$	χ^2/d.o.f.
All moments	$267.3^{+23.4}_{-19.1}$	$0.1226^{+0.0012}_{-0.0017}$	$752.665/(1847-(54+1))$
Even moments	$236.4^{+33.2}_{-54.9}$	$0.1203^{+0.0025}_{-0.0047}$	$334.339/(789-(27+1))$
Odd moments	$285.3^{+23.4}_{-23.5}$	$0.1239^{+0.0016}_{-0.0017}$	$415.28/(1058-(27+1))$

Table 7.1: In this table we present the results of the analysis performed with all moments included and with only odd or even moments included.

which is comparable to the value of 0.68 found in Ref. [193]. We also perform the analysis separately for even and odd moments only, and these results are presented together with the result for all moments in table 7.1.

Note that in the results in Eqs. (7.86) and (7.87) we have not included the theoretical error due to RS dependence and the unknown higher order corrections. However, due to its CORGI nature, this error is expected to be significantly lower in our analysis than that of the equivalent PS analysis of Refs. [192, 193].

In table 7.2 we present the values of the fitting parameters $A_n^{\rm NS}$ and $A_n^{\pm}$ corresponding to the CORGI fits in Figs. 7.1 and 7.2. We see that the A_n^+ parameters are all positive, corresponding to an overall negative contribution of the (+) component in Eq. (7.49). Using the values of the fitting parameters in table 7.2 together with Eq. (7.49) we can obtain values for the individual non-singlet, (+) and (−) components of the moments, as well as the combined singlet contribution, at any Q^2. In table 7.3 we present values for these components at $Q^2 = 14.9\,{\rm GeV}^2$. From these values we can obtain the determinants in Eq. (7.77). We find that most of these are positive, and those that are negative are extremely small in magnitude ($< 10^{-9}$).

We find that applying the constraints in Eqs. (7.64) - (7.67) and (7.71) is crucial to obtaining a physical minimum in χ^2 (by which we mean a minimum which corresponds to an acceptable value of $\Lambda_{\overline{\rm MS}}$). In contrast, the determinantal constraints are of less importance, in the sense that implementing them has little effect on the resultant value of $\Lambda_{\overline{\rm MS}}$. However, without these constraints we find that the values of the fitting parameters resulting from the fit are such that a substantial number of the determinantal constraints are violated.

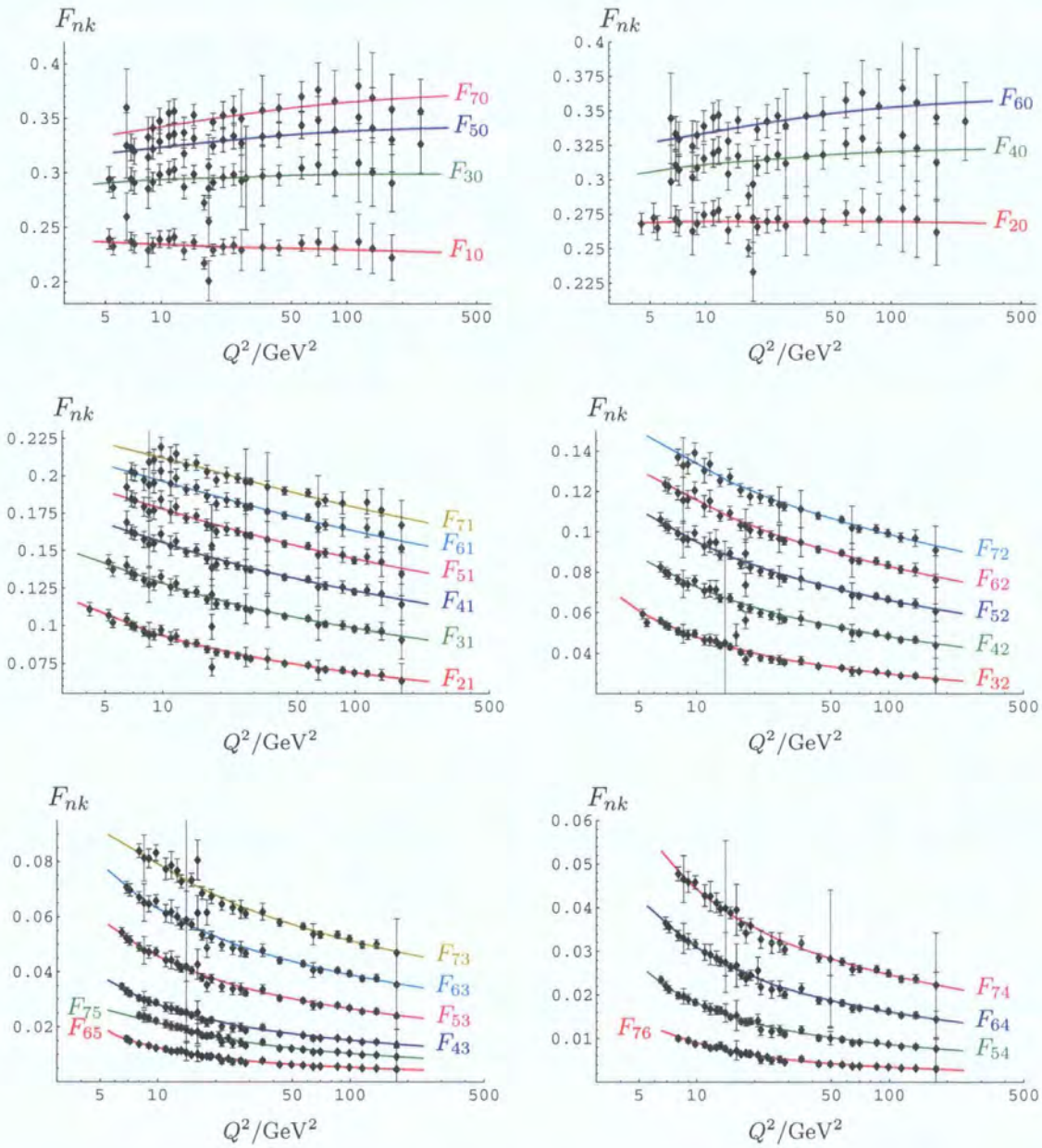


Figure 7.1: CORGI fits for the Bernstein averages of F_2, with TMCs included.

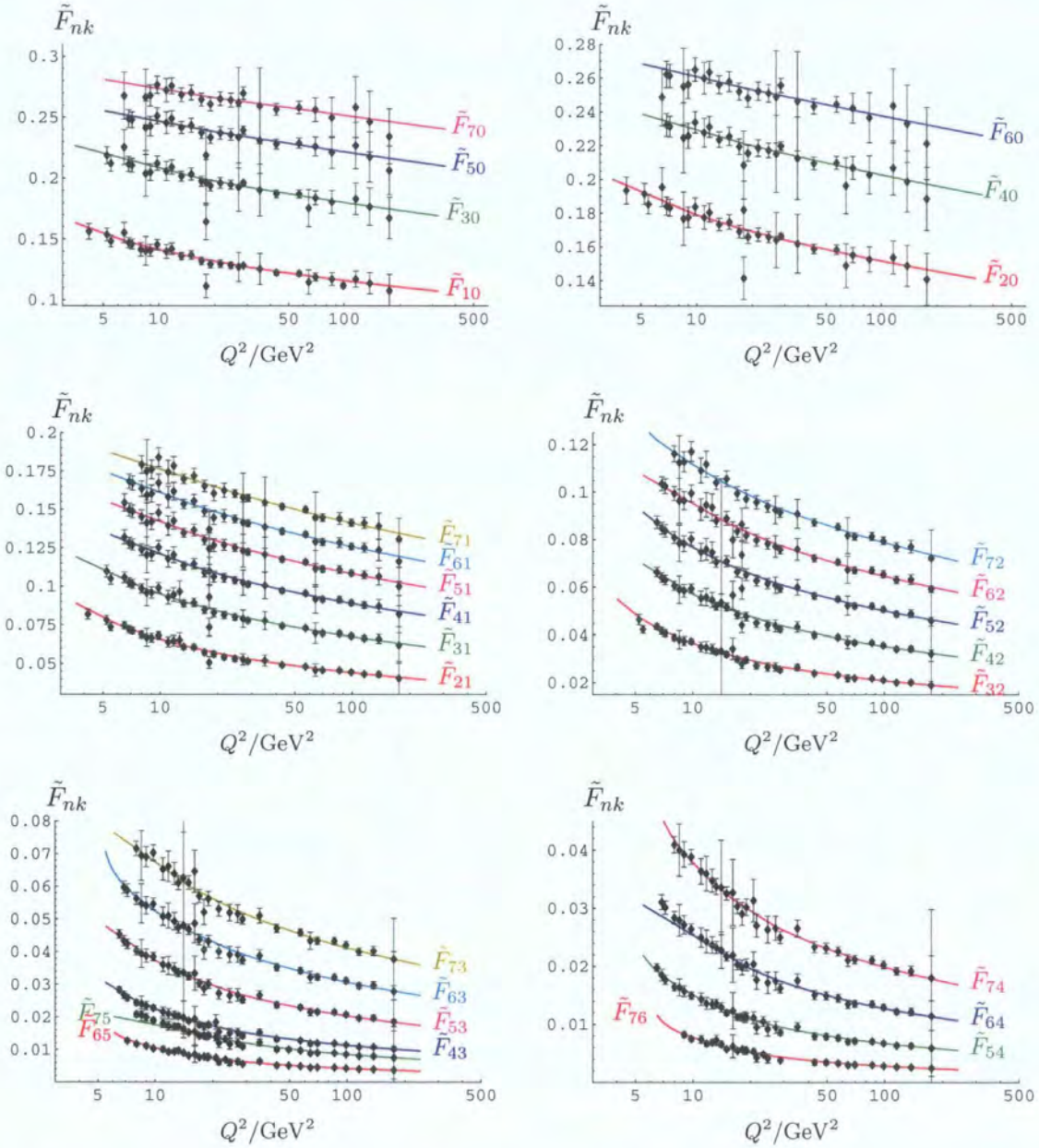


Figure 7.2: CORGI fits for the modified Bernstein averages of F_2, with TMCs included.

n	$A_n^{\mathrm{NS},(4)}$	$A_n^{-,(4)}$	$A_n^{+,(4)}$
2	0.3108	4.193×10^{-2}	0.2451
3	0.1265	2.271×10^{-2}	0.1414
4	6.719×10^{-2}	8.342×10^{-3}	7.488×10^{-2}
5	4.008×10^{-2}	3.569×10^{-3}	5.609×10^{-2}
6	2.562×10^{-2}	2.273×10^{-3}	4.655×10^{-2}
7	1.744×10^{-2}	1.212×10^{-3}	3.103×10^{-2}
8	1.22×10^{-2}	9.946×10^{-4}	2.924×10^{-2}
9	9.053×10^{-3}	5.446×10^{-4}	1.848×10^{-2}
10	6.693×10^{-3}	4.968×10^{-4}	1.824×10^{-2}
11	5.334×10^{-3}	2.71×10^{-4}	1.101×10^{-2}
12	4.089×10^{-3}	2.539×10^{-4}	1.083×10^{-2}
13	3.486×10^{-3}	1.437×10^{-4}	6.369×10^{-3}
14	2.726×10^{-3}	1.395×10^{-4}	6.486×10^{-3}
15	2.481×10^{-3}	9.303×10^{-5}	4.571×10^{-3}
16	1.966×10^{-3}	8.181×10^{-5}	4.066×10^{-3}
17	1.908×10^{-3}	7.334×10^{-5}	3.789×10^{-3}
18	1.5×10^{-3}	5.31×10^{-5}	2.745×10^{-3}
19	1.544×10^{-3}	6.35×10^{-5}	3.375×10^{-3}

Table 7.2: In the table above, we present the values of the fitting parameters A_n^{NS}, A_n^+ and A_n^- in the $N_f = 4$ region corresponding to the CORGI fits in Figs. (7.1) and (7.2).

The results in tables 7.1 - 7.3 and in Figs. 7.1 and 7.2 correspond to the analysis performed with TMCs included. When we remove these corrections from the analysis, we find a shift in the predicted value of $\Lambda_{\overline{\mathrm{MS}}}$ similar to that obtained in the F_3 analysis. For example, in the 'all moments' analysis without TMCs we obtain $\Lambda_{\overline{\mathrm{MS}}}^{(5)} = 306.0 \pm 17$ MeV with a χ^2 value of 795.3. Including HT corrections results in a smaller shift in the scale parameter of $\approx$ 10 MeV, and this is included in the error quoted in Eqs. (7.86) and (7.87). We also include a modelling error in these results, as we did for F_3.

We also perform a PS analysis of the F_2 moments in order to check consistency with previous analyses. A PS expression for F_2 moments can be assembled from the expression for the singlet component given in Eq. (7.39) and by modifying the non-singlet F_3 expression of Eq. (6.74) so that it applies to F_2. Unfortunately, when we perform the analysis using these expressions we find no stable minima in χ^2 for values of $\Lambda_{\overline{\mathrm{MS}}}^{(4)} < 590$ MeV. We discuss possible reasons for this in the next section.

n	$\mathcal{M}^{\rm NS}(n;Q_0^2)$	$\mathcal{M}^{-}(n;Q_0^2)$	$\mathcal{M}^{+}(n;Q_0^2)$	$\mathcal{M}^{\rm S}(n;Q_0^2)$
2	0.1696	6.338×10^{-2}	6.248×10^{-2}	8.997×10^{-4}
3	4.056×10^{-2}	8.43×10^{-3}	8.398×10^{-3}	3.2×10^{-5}
4	1.512×10^{-2}	1.622×10^{-3}	1.621×10^{-3}	1.138×10^{-6}
5	6.946×10^{-3}	6.239×10^{-4}	6.228×10^{-4}	1.103×10^{-6}
6	3.621×10^{-3}	3.237×10^{-4}	3.226×10^{-4}	1.095×10^{-6}
7	2.089×10^{-3}	1.461×10^{-4}	1.461×10^{-4}	1.163×10^{-8}
8	1.274×10^{-3}	1.044×10^{-4}	1.044×10^{-4}	9.363×10^{-10}
9	8.413×10^{-4}	5.084×10^{-5}	5.084×10^{-5}	9.363×10^{-10}
10	5.628×10^{-4}	4.196×10^{-5}	4.196×10^{-5}	9.363×10^{-10}
11	4.112×10^{-4}	2.099×10^{-5}	2.099×10^{-5}	9.363×10^{-10}
12	2.922×10^{-4}	1.829×10^{-5}	1.829×10^{-5}	9.363×10^{-10}
13	2.329×10^{-4}	9.35×10^{-6}	9.35×10^{-6}	4.7×10^{-15}
14	1.716×10^{-4}	8.754×10^{-6}	8.754×10^{-6}	4.694×10^{-15}
15	1.482×10^{-4}	5.548×10^{-6}	5.548×10^{-6}	4.685×10^{-15}
16	1.12×10^{-4}	4.657×10^{-6}	4.657×10^{-6}	2.195×10^{-15}
17	1.042×10^{-4}	4.003×10^{-6}	4.003×10^{-6}	2.195×10^{-15}
18	7.89×10^{-5}	2.791×10^{-6}	2.791×10^{-6}	2.421×10^{-20}
19	7.849×10^{-5}	3.227×10^{-6}	3.227×10^{-6}	1.991×10^{-20}

Table 7.3: In this table we give the values of the non-singlet, $(-)$, $(+)$ and singlet components of the moments resulting from the CORGI fits, evaluated at $Q_0^2 = 14.9$ GeV^2.

7.8 Summary and conclusions

In this chapter we attempted to expand the foundation of CORGI phenomenology by applying an analysis of the type carried out in chapter 6 to moments of F_2. We showed how NNLO expressions for the singlet component of the F_2 moments can be obtained and how this component can then be 'CORGI-ized'. We described how the Bernstein averages method can be adapted to F_2 and how the positivity constraints introduced in chapter 6 can be applied to the non-singlet and singlet components of F_2 individually.

The results of the CORGI analysis are presented in tables 7.1 - 7.2 and the fits to the data are shown in Figs. 7.1 and 7.2. The values of $\Lambda_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ corresponding to these fits are higher than the current global averages, but are still within acceptable proximity. There is also good consistency between odd and even sets of moments.

As was the case for the F_3 analysis, the value of χ^2/d.o.f. in Eq. (7.88) is rather lower than one might expect. However, this is consistent with the value found in

Refs. [192,193]. Furthermore, with regard to the correlation of the data for the averages, the same caveats apply to this analysis as to that of chapter 6; the true χ^2 minimum can only be obtained when the correlation is taken into account. Here we give a value of χ^2/d.o.f. corrected for the rescaling of errors applied in the χ^2 function. In analogy with Eq. (6.143),

$$\frac{\chi^2}{\text{d.o.f.}} = \frac{752.665}{270 + 347 - 55} = 1.34, \tag{7.89}$$

is more representative of the true χ^2 per d.o.f.

Target mass corrections appear to be as important for F_2 as they are for F_3. Omitting these corrections results in a $\approx +40$ MeV shift in the predicted value of $\Lambda_{\overline{\text{MS}}}$. Higher twist corrections also appear to be of equal importance to F_2 and F_3. The values of the singlet and non-singlet components of the moments at $Q^2 = 14.9\,\text{GeV}^2$ are given in table 7.3, and these values automatically satisfy the constraints in Eqs. (7.64) - (7.67) and (7.71). They also satisfy most of the determinantal constraints – only some of these constraints are violated and only by negligible magnitudes ($< 10^{-9}$).

For the PS analysis, we find there is no stable minimum in χ^2 that corresponds to a physical value of $\Lambda_{\overline{\text{MS}}}$. This is in contradiction to previous analyses [192, 193] in which stable minima were found for values of $\Lambda_{\overline{\text{MS}}}$ in good agreement with the global average. It is curious that the CORGI analysis succeeds where the PS analysis fails. Clearly we cannot claim this to be an endorsement of the CORGI approach over the PS approach, since previous PS analyses have been successful. Furthermore, the fact that the results of the non-singlet analysis in the previous chapter were perfectly consistent suggests that it is the singlet component which causes problems in the F_2 case.

An algebraic error would seem to be an obvious candidate for the reason the CORGI and PS analyses differ so drastically. However, the PS and CORGI coefficients are related by fairly simple relations, and having carried out checks on the algebra, we are confident that this is not the cause of the problem.

The main difference between the CORGI and physical scale expressions is the re-scaling of $\Lambda_{\overline{\text{MS}}}$ in the CORGI case via Eqs. (7.51) and (7.52). This re-scaling represents the resumming of an infinite set of terms which are RG-predictable at NLO, but the prac-

tical effect of this is to push the Landau pole in a_0 to higher scales as n increases (as can be seen from the table in section D.3). One can see how this might have the effect of favouring a lower value of $\Lambda_{\overline{\rm MS}}$ in the CORGI case, since pushing it too high would result in a_0 and therefore, the moments being of large magnitude at $Q^2 \sim 2\,\mathrm{GeV}^2$, making agreement with the data difficult. However, the PS result does contain a truncated set of the resummed terms which give rise to the re-scaling in the CORGI case (see Eq. (1.161)), and therefore one would expect these terms to mimic the re-scaling. Furthermore, the F_3 analysis showed no significant difference between the results of the PS and CORGI analysis, even though the re-scaling of $\Lambda_{\overline{\rm MS}}$ is of similar magnitude in the F_2 and F_3 cases.[3] This makes it difficult to argue that there should be some natural mechanism by which the CORGI and PS approaches might generate such drastically different results. Indeed, it is more likely to be some subtlety of the fitting procedure that we have not taken into account that is causing this impasse. Work is continuing on this analysis and we intend to report on it in a future publication.

[3]This can be seen by comparing the values of $\Lambda_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$ in the tables in sections C.1 and D.3 and noting that they are of similar magnitude.

Chapter 8

Summary and conclusions

In this thesis, we attempted to improve the predictive power of QCD by investigating how the large-order behaviour of perturbation theory affects the IR properties of observables, and also how the renormalization scheme dependence problem may be solved.

In chapter 1 we presented an introduction to perturbative QCD, focusing on the RS dependence of perturbative predictions. We introduced three proposed solutions to overcoming this problem: the PS, CORGI and EC approaches.

In chapter 2 we introduced the subject of deep inelastic lepton-hadron scattering. We described the parton model and how QCD corrections to it can be calculated. We then introduced structure functions and their Mellin moments, and described how the scale dependence of the moments can be determined by applying RG methods to composite operators. We also briefly described how the factorization process allows us to separate non-perturbative and perturbative effects and how, in analogy with renormalization, this leads to an FS dependence of factorized observables.

In chapter 3 (the last of the introductory chapters), we discussed the behaviour of perturbative QCD at large orders of the coupling. We described Dyson's original argument for the divergence of QED and detailed the two known sources of divergence in QCD: renormalons and instantons. We then studied renormalons in more detail and

described how their existence has inspired a series of all-orders leading-N_f/leading-b calculations for QCD observables. We then explained how, in the context of the Borel transform, we can separate the large-order influence of renormalons into UV and IR components and how it is the IR renormalons which cause ambiguities in the perturbative definition of an observable. We then discussed how such ambiguities can be linked to non-perturbative physics and how renormalon singularities in the Borel plane are in one-to-one correspondence with QCD condensates.

In chapter 4 we studied the IR properties of all-orders leading-b resummations for Euclidean observables. Our principle aim was to investigate the behaviour of these resummations at and below the Landau pole. From the singular behaviour of fixed-order predictions, and individual renormalon contributions at the Landau-pole, we would expect that the full all-orders expansion is singular too. Remarkably, this is not the case and cancellations between individual renormalons ensure that when we sum over the full set (or over special subsets) we recover a finite result. Although we might have expected the full perturbative expansion (written in terms of the all-orders coupling) to be finite at this point, the fact that a subset of this full set of terms (the leading-b components) sums to a finite result is very surprising. Indeed, this is the first example of a finite result being obtained at the Landau pole without some form of analytical continuation procedure being carried out in order to ensure finiteness.

We also found that below the Landau pole the all-orders predictions remain finite and freeze to zero in the $Q^2 \to 0$ limit. Hence, at $Q^2 = 0$ the value of the observable is simply the parton model result. We can understand this IR behaviour in terms of the skeleton expansion representation. When the one-loop coupling is used we find that the skeleton expansion separates naturally in to a UV region in which the Q^2-dependence can be obtained via the standard Borel representation, and an IR region in which the Q^2-dependence is given by means of a modified Borel representation. In the standard representation, ambiguities are generated by IR renormalons, however in the modified representation the singularity structure of the Borel transform is inverted and it is the UV renormalons that generate ambiguities. We evaluated the form of these ambiguities in the two regions, and by demanding that they are compensated by their non-perturbative counterparts, we were led to a new model for power corrections. This results in our being able to write down a single, simple expression for both the perturbative and non-perturbative components of an observable in terms of the QCD

scale parameter Λ and a single non-perturbative parameter κ.

In chapter 5 we applied the results of chapter 4 to a phenomenological analysis of the GLS, polarized Bjorken and unpolarized Bjorken sum rules. We adapted the leading-b resummations to CORGI perturbation theory and from this we obtained predictions written in terms of the 't Hooft coupling which resummed the leading-b components of the RS invariant quantities X_n to all orders. We compared these results with the NLO and NNLO CORGI predictions, in order to assess the validity of fixed-order perturbation theory. We found that the fixed-order and all-orders estimates differ drastically for the GLS and polarized Bjorken below $\sim 5\text{GeV}^2$ and for the unpolarized Bjorken below $\sim 2\text{GeV}^2$. We also compared the fixed-order and all-orders predictions with the available data. All predictions were in reasonably good agreement with experiment, but the sizeable error bars limit the significance of any conclusions we can draw from this comparison. We also tested the model for non-perturbative power corrections proposed in chapter 4 on these sum rules and found that the data favours power corrections of reasonably small magnitude.

In chapter 6 we turned our attention to the FRS dependence of fixed order predictions for moments of structure functions, performing a phenomenological analysis of F_3 moments in neutrino-nucleon scattering. We adopted the three different approaches to solving the RS dependence problem that were outlined in chapter 1. We showed how these approaches are generalized to deal with FRS dependence and we used the method of Bernstein averages in order to overcome the problems associated with regions of x and Q^2 for which structure function data is unavailable. We also improved upon previous analyses of this kind by including both odd and even moments and also by including higher moments. We found that the CORGI, PS and EC approaches all predicted similar values of the QCD scale parameter and that these were in excellent agreement with the current global average (see tables 6.1 and 6.3). The errors on these values were greatly reduced in comparison with previous analyses. We also found consistency between analyses in which we included only odd or even moments. We attempted to quantify the success of the Bernstein average method by measuring the deviation in the final result generated by changing the method used to obtain the experimental averages and found that this deviation is reasonably small. We also found that whereas target mass corrections appear to be significant for F_3, higher twist corrections are less so.

In chapter 7 we performed a phenomenological analysis of moments of F_2. As for the F_3 analysis, we used the method of Bernstein averages, whilst recent results also allowed us to extend the set of moments used relative to previous analyses. In the case of F_2, matters are complicated by the presence of a singlet contribution. We showed how an NNLO expression for this component can be derived and how we can then apply CORGI perturbation theory to it. From the CORGI analysis we found that the F_2 data favours a value of the QCD scale parameter which is in reasonably good agreement with the current global average (see table 7.1). We found good consistency between odd and even moments and also that TMCs and HT effects have a similar impact on the outcome of the analysis as they did in the F_3 case, shifting the predicted value of $\Lambda_{\overline{\rm MS}}$ by ≈ 40 and ≈ 10 MeV respectively. In the PS analysis we found that a value of the QCD scale parameter of $\Lambda^{(4)}_{\overline{\rm MS}} > 590$ MeV was favoured, which is clearly unphysical. We believe that this is due to subtleties of the fitting procedure related to the presence of the singlet contribution (the F_2 analysis has triple the number of fitting parameters of the F_3 analysis). It is curious that the CORGI approach is successful where the PS approach fails. However we cannot claim that this is an endorsement of CORGI over PS since previous PS analyses have found perfectly physically valid values of $\Lambda_{\overline{\rm MS}}$. Work is continuing on this analysis.

Appendix A

Group theory relations for SU(N_C)

A matrix representation of the generators of SU(3) is provided by the eight Gell-Mann matrices, λ_a. These matrices can be found in many textbooks, including Ref. [3]. By convention, the normalization of these matrices is chosen such that,

$$\mathrm{Tr}(T^a T^b) = T_R \delta^{ab}, \qquad \sum_a T^a_{ij} T^a_{jk} = C_F \delta_{jk}. \tag{A.1}$$

As a result, the matrices obey the following relations,

$$\{T^a, T^b\} = \frac{1}{N_c}\delta^{ab} I + d^{abc} T^c, \tag{A.2}$$

$$\sum_{a,b} d^{abc} d^{abd} = \frac{N_c^2 - 4}{N_c}\delta^{cd}, \qquad d^{aad} \equiv 0, \tag{A.3}$$

$$(T^a T^b T^a)_{ij} = T^b_{ij}\left(C_F - \frac{C_A}{2}\right). \tag{A.4}$$

C_A and C_F are group theory factors. Their values, in SU(N_c) are,

$$C_A = N_c, \qquad C_F = \frac{N_c^2 - 1}{2N_c}. \tag{A.5}$$

Appendix B

Special functions

In this appendix we define some of the special functions used in this thesis.

B.1 The Gamma function

The Gamma function is defined as,

$$\Gamma(z) \equiv \int_0^\infty t^{z-1}\mathrm{e}^{-t}\,dt. \tag{B.1}$$

For integer n, it satisfies

$$n! = \Gamma(n+1), \tag{B.2}$$

and is therefore a generalization of the factorial function to non-integer values of n.

The Gamma function can be used to determine the large n behaviour of the factorial function via Stirling's formula:

$$n! \underset{n\to\infty}{\approx} \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n. \tag{B.3}$$

For negative integer values of z, and also at $z = 0$, the Gamma function is singular. In the small z limit we have,

$$\Gamma(\varepsilon) = \frac{1}{\varepsilon} - \gamma_{\mathrm{E}} + \mathcal{O}(\varepsilon), \tag{B.4}$$

where $\gamma_{\rm E}$ is the Euler-Mascheroni constant, and $\gamma_{\rm E} = 0.577215665\ldots$.

B.2 The beta function

The beta function (not to be confused with the beta-function equation) is defined in terms of gamma functions,

$$\text{Beta}(x,y) \equiv \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \tag{B.5}$$

B.3 The exponential integral function

The exponential integral function is defined as,

$$\text{Ei}(x) \equiv -\int_{-x}^{\infty} dt\, \frac{\mathrm{e}^{-t}}{t}, \tag{B.6}$$

for $x < 0$, and can be defined for $x > 0$ by taking the principal value of the above integral.

B.4 The Riemann zeta function

The Riemann zeta function is defined as,

$$\zeta(s) \equiv \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{B.7}$$

This sum converges for all s with $\text{Re}(s) > 1$, and for the first few integer values of s has the following values,

$$\zeta(2) = \frac{\pi^2}{6}, \tag{B.8}$$

$$\zeta(3) = 1.20206\ldots, \tag{B.9}$$

$$\zeta(4) = \frac{\pi^4}{90}, \tag{B.10}$$

$$\zeta(5) = 1.03693\ldots, \tag{B.11}$$

$$\zeta(6) = \frac{\pi^6}{945}. \tag{B.12}$$

B.5 The dilogarithm function

The dilogarithm function can be defined via the integral,

$$\mathrm{Li}_2(z) \equiv \int_z^0 \frac{\ln(1-t)}{t}\, dt, \tag{B.13}$$

or by the sum,

$$\mathrm{Li}_2(z) \equiv \sum_{n=1}^{\infty} \frac{z^n}{n^2}. \tag{B.14}$$

At the point $z = 1$, this is simply equal to $\zeta(2)$.

Appendix C

NLO and NNLO coefficients for moments of F_3 structure functions

In this appendix we present the coefficients for the NNLO predictions of F_3 moments in CORGI, PS and EC perturbation theory. These predictions are given in Eqs. (6.65), (6.74) and (6.81).

C.1 CORGI coefficients

In the table below, we present the information required to construct an NNLO CORGI prediction for the moments of F_3. The coefficient d/b is relevant to all three approaches, $\Lambda_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$ represents the NLO contribution to Eq. (6.65) (or Eq. (6.72) in the case of $n = 1$), and X_2 is the NNLO coefficient.

	$N_f = 4$			$N_f = 5$		
n	d/b	$\Lambda_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	X_2	d/b	$\Lambda_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	X_2
1	0.	2.439	n/a	0.	2.456	n/a
2	0.4267	1.578	−2.448	0.4638	1.609	−3.01
3	0.6667	2.239	−2.203	0.7246	2.269	−3.28
4	0.8373	2.655	−2.059	0.9101	2.68	−3.485
5	0.9707	2.977	−1.933	1.055	3.	−3.627
6	1.08	3.242	−1.797	1.174	3.262	−3.713
7	1.174	3.472	−1.661	1.276	3.489	−3.766
8	1.255	3.674	−1.521	1.364	3.69	−3.792
9	1.327	3.857	−1.381	1.442	3.871	−3.8
10	1.392	4.023	−1.24	1.513	4.036	−3.793
11	1.45	4.177	−1.1	1.576	4.188	−3.776
12	1.504	4.32	−0.9598	1.635	4.33	−3.749
13	1.554	4.454	−0.8213	1.689	4.463	−3.716
14	1.599	4.579	−0.6841	1.739	4.587	−3.677
15	1.642	4.698	−0.5483	1.785	4.705	−3.633
16	1.682	4.811	−0.4142	1.829	4.817	−3.586
17	1.72	4.918	−0.2816	1.87	4.924	−3.536
18	1.756	5.021	−0.1507	1.909	5.026	−3.483
19	1.79	5.119	−0.02136	1.945	5.123	−3.428
20	1.822	5.213	0.1062	1.98	5.217	−3.372

C.2 Physical scale coefficients

In the table below, we present the information required to construct an NNLO PS prediction for the moments of F_3. R_1 and R_2 are the NLO and NNLO coefficients respectively in Eq. (6.74).

	$N_f = 4$		$N_f = 5$	
n	R_1	R_2	R_1	R_2
1	$-1.$	-3.25	$-1.$	-2.917
2	0.6119	-3.123	0.6006	-3.125
3	1.928	0.4136	1.893	-0.08583
4	3.017	5.373	2.959	4.363
5	3.961	11.03	3.886	9.523
6	4.792	17.02	4.702	15.03
7	5.541	23.19	5.438	20.73
8	6.221	29.42	6.107	26.51
9	6.846	35.67	6.722	32.33
10	7.424	41.88	7.292	38.13
11	7.964	48.06	7.823	43.91
12	8.47	54.17	8.321	49.64
13	8.946	60.22	8.79	55.31
14	9.396	66.18	9.234	60.92
15	9.823	72.07	9.655	66.46
16	10.23	77.88	10.06	71.93
17	10.62	83.61	10.44	77.33
18	10.99	89.26	10.81	82.66
19	11.34	94.83	11.16	87.93
20	11.69	100.3	11.49	93.12

C.3 Effective charge coefficients

In the table below, we present the information required to construct an NNLO EC prediction for the moments of F_3.

	$N_f = 4$		$N_f = 5$	
n	$\Lambda^{\rm EC}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	$\tilde{X}_2$	$\Lambda^{\rm EC}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	$\tilde{X}_2$
1	2.439	n/a	2.456	n/a
2	1.09	−3.366	1.158	−4.901
3	1.547	−0.9329	1.633	−2.937
4	1.834	−0.08705	1.929	−2.239
5	2.057	0.3803	2.159	−1.848
6	2.24	0.7083	2.348	−1.572
7	2.399	0.9563	2.511	−1.362
8	2.539	1.16	2.656	−1.19
9	2.665	1.331	2.786	−1.045
10	2.78	1.481	2.905	−0.918
11	2.886	1.613	3.014	−0.8054
12	2.985	1.733	3.116	−0.7037
13	3.077	1.843	3.212	−0.6108
14	3.164	1.944	3.302	−0.5251
15	3.246	2.038	3.386	−0.4456
16	3.324	2.125	3.467	−0.3712
17	3.399	2.208	3.544	−0.3012
18	3.47	2.286	3.617	−0.2352
19	3.537	2.36	3.687	−0.1726
20	3.603	2.43	3.755	−0.113

Appendix D

LO, NLO and NNLO coefficients for moments of F_2

D.1 Differences in notation between this thesis and other works

There are several differences in convention between the way we define the anomalous dimension and the ways it is defined by Yndurain (in Ref. [192]) and by Buras (in Ref. [207]). Firstly, the definition of the anomalous dimension itself differs. We have,

$$\gamma_{\mathrm{S},n}^{[\mathrm{Us}]}(a) \;=\; -\hat{\gamma}_{[\mathrm{Bur.}]}^{n}(g) \;=\; \gamma_{[\mathrm{Ynd.}]}(a). \tag{D.1}$$

Also, due to our definition of a, the individual coefficients of the anomalous dimension differ:

$$4^{i+1}\boldsymbol{d}_i^{[\mathrm{Us}]} \;=\; \gamma_{[\mathrm{Ynd.}]}^{(i)} \;=\; \gamma_{[\mathrm{Bur.}]}^{(i),n}. \tag{D.2}$$

Finally, the eigenvalues of the first coefficient matrix of the anomalous dimension are different.

$$\lambda_{\pm}^{[\mathrm{Us}]} \;=\; b\,d_{\pm}^{[\mathrm{Bur.}]} \;=\; -b\,d_{\mp}^{[\mathrm{Ynd.}]}. \tag{D.3}$$

D.2 LO coefficients

Here, we present the leading order coefficients relevant to the non-singlet component, and the (+) and (−) singlet components of the F_2 moments.

	$N_f = 4$			$N_f = 5$		
n	$d^{\rm NS}(n)/b$	λ_-/b	λ_+/b	$d^{\rm NS}(n)/b$	λ_-/b	λ_+/b
2	0.4267	0.	0.7467	0.4638	0.	0.8986
3	0.6667	0.6085	1.386	0.7246	0.6542	1.601
4	0.8373	0.817	1.852	0.9101	0.8846	2.104
5	0.9707	0.9604	2.192	1.055	1.042	2.471
6	1.08	1.074	2.46	1.174	1.166	2.762
7	1.174	1.17	2.683	1.276	1.271	3.005
8	1.255	1.252	2.875	1.364	1.36	3.212
9	1.327	1.325	3.043	1.442	1.44	3.395
10	1.392	1.39	3.192	1.513	1.511	3.557
11	1.45	1.449	3.328	1.576	1.575	3.704
12	1.504	1.503	3.451	1.635	1.633	3.838
13	1.554	1.553	3.564	1.689	1.687	3.961
14	1.599	1.599	3.669	1.739	1.738	4.075
15	1.642	1.642	3.767	1.785	1.784	4.181
16	1.682	1.682	3.858	1.829	1.828	4.281
17	1.72	1.72	3.944	1.87	1.869	4.374
18	1.756	1.756	4.025	1.909	1.908	4.462
19	1.79	1.789	4.101	1.945	1.945	4.545
20	1.822	1.821	4.174	1.98	1.98	4.624

D.3 NLO CORGI coefficients

In the table below we present the NLO information relevant to the CORGI predictions for the non-singlet, $(-)$ and $(+)$ components of F_2 moments, given in Eqs. (7.49) - (7.52).

	$N_f = 4$			$N_f = 5$		
n	$\Lambda^{\rm NS}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	$\Lambda^{-}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	$\Lambda^{+}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	$\Lambda^{\rm NS}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	$\Lambda^{-}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$	$\Lambda^{+}_{\mathcal{M}}/\Lambda_{\overline{\rm MS}}$
2	2.156	1.804	2.38	2.199	1.909	2.517
3	2.575	2.671	2.819	2.609	3.731	2.886
4	2.893	1.698	2.903	2.921	2.565	2.922
5	3.163	3.05	3.313	3.187	3.091	3.326
6	3.394	3.343	3.501	3.415	3.364	3.502
7	3.601	3.576	3.806	3.619	3.592	3.805
8	3.786	3.771	3.982	3.803	3.785	3.975
9	3.956	3.946	4.223	3.971	3.959	4.214
10	4.113	4.106	4.38	4.126	4.117	4.366
11	4.258	4.253	4.579	4.27	4.264	4.564
12	4.394	4.39	4.719	4.404	4.4	4.701
13	4.522	4.519	4.89	4.531	4.528	4.87
14	4.643	4.641	5.016	4.651	4.648	4.994
15	4.758	4.756	5.165	4.765	4.763	5.142
16	4.867	4.865	5.28	4.873	4.871	5.256
17	4.971	4.97	5.413	4.977	4.975	5.387
18	5.071	5.07	5.519	5.076	5.074	5.492
19	5.166	5.166	5.639	5.171	5.17	5.611
20	5.258	5.258	5.737	5.262	5.261	5.707

D.4 NNLO CORGI coefficients

In the table below we present the NNLO coefficients of the CORGI predictions for the non-singlet, $(-)$ and $(+)$ components of F_2 moments, given in Eqs. (7.49) - (7.52).

	$N_f = 4$			$N_f = 5$		
n	$X_2^{\rm NS}$	X_2^-	X_2^+	$X_2^{\rm NS}$	X_2^-	X_2^+
2	-0.7744	-1.991	-2.485	-1.557	-4.398	-41.32
3	-1.145	-1.075	-5.596	-2.378	-8.502	-3.738
4	-1.274	-2.126	-3.547	-2.823	-6.93	-2.363
5	-1.312	-1.094	-5.608	-3.109	-9.611	-2.875
6	-1.284	-1.136	-3.13	-3.289	-7.468	-3.067
7	-1.228	-1.128	-4.305	-3.411	-9.062	-3.199
8	-1.146	-1.055	-2.07	-3.487	-7.079	-3.195
9	-1.053	-0.9668	-2.807	-3.536	-8.135	-2.798
10	-0.9479	-0.8514	-0.884	-3.56	-6.418	-4.141
11	-0.8384	-0.7128	-1.368	-3.568	-7.161	-3.743
12	-0.7234	-0.4785	0.2864	-3.563	-5.684	-3.658
13	-0.6065	-1.549	-0.0352	-3.548	-6.224	-3.61
14	-0.487	-0.621	1.403	-3.524	-4.942	-3.568
15	-0.3672	-0.4336	1.192	-3.494	-5.343	-3.527
16	-0.2464	-0.2884	2.458	-3.457	-4.216	-3.484
17	-0.126	-0.1557	2.326	-3.417	-4.517	-3.439
18	-0.005423	-0.02787	3.454	-3.373	-3.516	-3.391
19	0.1143	0.09657	3.38	-3.326	-3.742	-3.342
20	0.2337	0.2193	4.395	-3.277	-2.843	-3.29

D.5 NLO physical scale coefficients

In the table below we present the NLO coefficients of the PS predictions for the non-singlet, $(-)$ and $(+)$ components of F_2 moments, given in Eqs. (6.74) - (7.39).

	$N_f = 4$			$N_f = 5$		
n	$R_1^{\rm NS}$	R_1^-	R_1^+	$R_1^{\rm NS}$	R_1^-	R_1^+
2	1.167	-2.029	2.35	1.156	2.705	-6.206
3	2.317	2.208	5.341	2.282	5.659	2.956
4	3.317	1.423	7.364	3.259	7.537	2.727
5	4.205	4.015	9.919	4.13	10.08	3.957
6	4.998	4.902	11.7	4.909	11.81	4.808
7	5.719	5.666	13.7	5.617	13.8	5.557
8	6.378	6.342	15.21	6.264	15.3	6.222
9	6.987	6.962	16.85	6.863	16.92	6.833
10	7.552	7.533	18.16	7.419	18.22	7.396
11	8.08	8.066	19.55	7.939	19.6	7.922
12	8.576	8.565	20.7	8.428	20.75	8.414
13	9.045	9.036	21.91	8.889	21.95	8.878
14	9.488	9.481	22.95	9.326	22.97	9.317
15	9.909	9.903	24.02	9.741	24.04	9.734
16	10.31	10.31	24.95	10.14	24.97	10.13
17	10.69	10.69	25.92	10.52	25.93	10.51
18	11.06	11.06	26.77	10.88	26.78	10.87
19	11.41	11.41	27.65	11.22	27.65	11.22
20	11.75	11.75	28.44	11.56	28.43	11.56

D.6 NNLO physical scale coefficients

In the table below we present the NNLO coefficients of the PS predictions for the non-singlet, $(-)$ and $(+)$ components of F_2 moments, given in Eqs. (6.74) - (7.39).

	$N_f = 4$			$N_f = 5$		
n	$R_2^{\rm NS}$	R_2^-	R_2^+	$R_2^{\rm NS}$	R_2^-	R_2^+
2	0.2038	-4.045	1.698	-0.1318	2.004	-12.11
3	3.535	3.514	14.74	2.748	15.15	6.348
4	8.243	-2.365	32.55	6.983	31.87	4.253
5	13.68	12.43	59.35	11.95	58.09	10.93
6	19.48	18.79	85.64	17.29	83.51	16.68
7	25.48	25.08	116.3	22.84	113.5	22.52
8	31.58	31.3	145.2	28.5	141.6	28.39
9	37.7	37.52	176.5	34.21	172.3	34.65
10	43.81	43.7	206.	39.92	201.	39.09
11	49.9	49.85	237.	45.61	231.4	45.24
12	55.92	56.03	266.2	51.27	259.9	51.
13	61.89	60.83	296.5	56.87	289.6	56.67
14	67.79	67.56	325.2	62.42	317.7	62.25
15	73.62	73.47	354.7	67.9	346.5	67.76
16	79.37	79.25	382.7	73.32	374.	73.2
17	85.05	84.95	411.3	78.68	402.	78.57
18	90.65	90.57	438.7	83.96	428.7	83.87
19	96.18	96.11	466.4	89.19	455.9	89.11
20	101.6	101.6	493.	94.34	481.9	94.27

References

[1] T. Muta, *Foundations of Quantum Chromodynamics: An Introduction to Perturbative Methods in Gauge Theories* (World Scientific, 1998).

[2] D. Bailin and A. Love, *Introduction to Gauge Field Theory* (Institute of Physics Publishing, 1993).

[3] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, 1996).

[4] F. J. Yndurain, *Quantum Chromodynamics: An Introduction to the Theory of Quarks and Gluons* (Springer-Verlag, 1983).

[5] R. Brock *et al.*, 'Handbook of perturbative QCD; version 1.1', (1994), Fermilab-Pub-94-316.

[6] W. Greiner and A. Schäfer, *Quantum Chromodynamics* (Springer, 1994).

[7] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* (Wiley, 1984).

[8] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Field Theory.* (McGraw-Hill, 1967).

[9] G. Sterman, *An Introduction to Quantum Field Theory* (Cambridge University Press, 1993).

[10] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. I. Troian, *Basics of perturbative QCD* (World Scientific, 1991).

[11] C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, 1985).

[12] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, 1995).

[13] E. S. Abers and B. W. Lee, 'Gauge theories', *Phys. Rept.* **9**, 1 (1973).

[14] G. Leibbrandt, 'Introduction to noncovariant gauges', *Rev. Mod. Phys.* **59**, 1067 (1987).

[15] A. Bassetto, G. Nardelli, and R. Soldati, *Yang-Mills Theories in Algebraic Noncovariant Gauges: Canonical Quantization and Renormalization* (World Scientific, 1991).

[16] L. D. Faddeev and V. N. Popov, 'Feynman diagrams for the Yang-Mills field', *Phys. Lett.* **B25**, 29 (1967).

[17] C. Becchi, A. Rouet, and R. Stora, 'Renormalization of gauge theories', *Annals Phys.* **98**, 287 (1976).

[18] M. Z. Iofa and I. V. Tyutin, 'Gauge invariance of spontaneously broken non-abelian theories in the Bogolyubov-Parasiuk-Hepp-Zimmerman method. (in russian)', *Teor. Mat. Fiz.* **27**, 38 (1976).

[19] P. Dirac, 'The Lagrangian in quantum mechanics', *Phys. Zeit der Sowjet Union* **3**, 64 (1933).

[20] R. P. Feynman, 'Space-time approach to nonrelativistic quantum mechanics', *Rev. Mod. Phys.* **20**, 367 (1948).

[21] S. A. Pernice and G. Oleaga, 'On the divergence of perturbation theory: Steps towards a convergent series', *Phys. Rev.* **D57**, 1144 (1998), `hep-th/9609139`.

[22] Y. Meurice, 'A simple method to make asymptotic series of Feynman diagrams converge', *Phys. Rev. Lett.* **88**, 141601 (2002), `hep-th/0103134`.

[23] Y. Meurice, 'Large field cutoffs make perturbative series converge', *Nucl. Phys. Proc. Suppl.* **106**, 908 (2002), `hep-lat/0110146`.

[24] G. 't Hooft and M. J. G. Veltman, 'Diagrammar', In *Particle Interactions At Very High Energies, Part B*, proceedings of the Summer Institute, Louvain 1973, p. 177 (Plenum, 1973), also in CERN-73-9.

[25] D. J. Gross and F. Wilczek, 'Ultraviolet behavior of non-Abelian gauge theories', *Phys. Rev. Lett.* **30**, 1343 (1973).

[26] H. D. Politzer, 'Reliable perturbative results for strong interactions?', *Phys. Rev. Lett.* **30**, 1346 (1973).

[27] W. E. Caswell, 'Asymptotic behavior of nonabelian gauge theories to two loop order', *Phys. Rev. Lett.* **33**, 244 (1974).

[28] D. R. T. Jones, 'Two loop diagrams in Yang-Mills theory', *Nucl. Phys.* **B75**, 531 (1974).

[29] O. V. Tarasov, A. A. Vladimirov, and A. Y. Zharkov, 'The Gell-Mann-Low function of QCD in the three loop approximation', *Phys. Lett.* **B93**, 429 (1980).

[30] S. A. Larin and J. A. M. Vermaseren, 'The three loop QCD beta function and anomalous dimensions', *Phys. Lett.* **B303**, 334 (1993), `hep-ph/9302208`.

[31] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, 'The four-loop beta function in quantum chromodynamics', *Phys. Lett.* **B400**, 379 (1997), `hep-ph/9701390`.

[32] W.-M. Y. *et al.*, 'Review of particle physics', *Journal of Physics G* **33**, 1 (2006).

[33] P. M. Stevenson, 'Optimized perturbation theory', *Phys. Rev.* **D23**, 2916 (1981).

[34] R. Corless, *et al.*, 'On the lambert W function', *Advances in Computational Mathematics* **5**, 329 (1996), `http://www.apmaths.uwo.ca/~djeffrey/offprints.html`.

[35] E. Gardi, G. Grunberg, and M. Karliner, 'Can the QCD running coupling have a causal analyticity structure?', *JHEP* **07**, 007 (1998), `hep-ph/9806462`.

[36] B. A. Magradze, 'Analytic approach to perturbative QCD', *Int. J. Mod. Phys.* **A15**, 2715 (2000), `hep-ph/9911456`.

[37] D. T. Barclay, C. J. Maxwell, and M. T. Reader, 'Extracting $\Lambda_{\overline{\rm MS}}$ from experiment', *Phys. Rev.* **D49**, 3480 (1994).

[38] C. J. Maxwell, 'Complete renormalization group improvement: Avoiding scale dependence in QCD predictions', (1999), `hep-ph/9908463`.

[39] C. J. Maxwell and A. Mirjalili, 'Complete renormalization group improvement: Avoiding factorization and renormalization scale dependence in QCD predictions', *Nucl. Phys.* **B577**, 209 (2000), `hep-ph/0002204`.

[40] G. 't Hooft: *The Whys of Subnuclear Physics*, Proceedings of the 15^{th} International School of Subnuclear Physics, Enrice, Italy (Plenum, 1977).

[41] G. Grunberg, 'Renormalization scheme independent QCD and QED: The method of effective charges', *Phys. Rev.* **D29**, 2315 (1984).

[42] G. Grunberg, 'Renormalization group improved perturbative QCD', *Phys. Lett.* **B95**, 70 (1980).

[43] K. G. Wilson, 'Non-lagrangian models of current algebra', *Phys. Rev.* **179**, 1499 (1969).

[44] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, 'QCD and resonance physics. Applications', *Nucl. Phys.* **B147**, 448 (1979).

[45] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, 'QCD and resonance physics. Sum rules', *Nucl. Phys.* **B147**, 385 (1979).

[46] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, 'QCD and resonance physics. The ρ - ω mixing', *Nucl. Phys.* **B147**, 519 (1979).

[47] S. Eidelman, F. Jegerlehner, A. L. Kataev, and O. Veretin, 'Testing nonperturbative strong interaction effects via the Adler function', *Phys. Lett.* **B454**, 369 (1999), `hep-ph/9812521`.

[48] D. M. Howe and C. J. Maxwell, 'All-orders infrared freezing of observables in perturbative QCD', *Phys. Rev.* **D70**, 014002 (2004), `hep-ph/0303163`.

[49] K. G. Chetyrkin, A. L. Kataev, and F. V. Tkachov, 'Higher order corrections to $\sigma_{tot}(e^+e^- \to \text{hadrons})$ in quantum chromodynamics', *Phys. Lett.* **B85**, 277 (1979).

[50] M. Dine and J. R. Sapirstein, 'Higher order QCD corrections in e^+e^- annihilation', *Phys. Rev. Lett.* **43**, 668 (1979).

[51] S. G. Gorishnii, A. L. Kataev, and S. A. Larin, 'The $\mathcal{O}(\alpha_s^3)$ corrections to $\sigma_{tot}(e^+e^- \to \text{hadrons})$ and $\Gamma(\tau \to \nu_\tau + \text{hadrons})$ in QCD', *Phys. Lett.* **B259**, 144 (1991).

[52] L. R. Surguladze and M. A. Samuel, 'Total hadronic cross-section in e^+e^- annihilation at the four loop level of perturbative QCD', *Phys. Rev. Lett.* **66**, 560 (1991).

[53] W. Celmaster and R. J. Gonsalves, 'An analytic calculation of higher order quantum chromodynamic corrections in e^+e^- annihilation', *Phys. Rev. Lett.* **44**, 560 (1980).

[54] R. D. Field, *Applications of Perturbative QCD* (Addison-Wesley, 1989).

[55] A. H. Mueller, *Perturbative Quantum Chromodynamics* (World Scientific, 1989).

[56] R. P. Feynman, 'Very high-energy collisions of hadrons', *Phys. Rev. Lett.* **23**, 1415 (1969).

[57] J. D. Bjorken and E. A. Paschos, 'Inelastic electron proton and gamma proton scattering, and the structure of the nucleon', *Phys. Rev.* **185**, 1975 (1969).

[58] R. P. Feynman, *Photon-hadron interactions* (Reading, 1972).

[59] R. G. Roberts, *The Structure of the proton: Deep inelastic scattering* (Cambridge University Press, 1990).

[60] D. J. Gross and F. Wilczek, 'Asymptotically free gauge theories. 2', *Phys. Rev.* **D9**, 980 (1974).

[61] H. Georgi and H. D. Politzer, 'Electroproduction scaling in an asymptotically free theory of strong interactions', *Phys. Rev.* **D9**, 416 (1974).

[62] H. D. Politzer, 'Asymptotic freedom: An approach to strong interactions', *Phys. Rept.* **14**, 129 (1974).

[63] A. Peterman, 'Renormalization group and the deep structure of the proton', *Phys. Rept.* **53**, 157 (1979).

[64] B. Geyer, D. Robaschik, and E. Wieczorek, 'Theory of deep inelastic lepton - hadron scattering. 1', *Fortschr. Phys.* **27**, 75 (1979).

[65] D. J. Gross and F. Wilczek, 'Asymptotically free gauge theories. 1', *Phys. Rev.* **D8**, 3633 (1973).

[66] E. G. Floratos, D. A. Ross, and C. T. Sachrajda, 'Higher order effects in asymptotically free gauge theories: The anomalous dimensions of Wilson operators', *Nucl. Phys.* **B129**, 66 (1977).

[67] 'Erratum', *Nucl. Phys.* **B139**, 545 (1978).

[68] E. G. Floratos, D. A. Ross, and C. T. Sachrajda, 'Higher order effects in asymptotically free gauge theories. 2. Flavor singlet wilson operators and coefficient functions', *Nucl. Phys.* **B152**, 493 (1979).

[69] A. Gonzalez-Arroyo and C. Lopez, 'Second order contributions to the structure functions in deep inelastic scattering. 3. The singlet case', *Nucl. Phys.* **B166**, 429 (1980).

[70] A. Gonzalez-Arroyo, C. Lopez, and F. J. Yndurain, 'Second order contributions to the structure functions in deep inelastic scattering. I. Theoretical calculations', *Nucl. Phys.* **B153**, 161 (1979).

[71] W. Furmanski and R. Petronzio, 'Singlet parton densities beyond leading order', *Phys. Lett.* **B97**, 437 (1980).

[72] G. Curci, W. Furmanski, and R. Petronzio, 'Evolution of parton densities beyond leading order: The nonsinglet case', *Nucl. Phys.* **B175**, 27 (1980).

[73] C. Lopez and F. J. Yndurain, 'Behavior at $x = 0$, 1, sum rules and parametrizations for structure functions beyond the leading order', *Nucl. Phys.* **B183**, 157 (1981).

[74] R. Hamberg and W. L. van Neerven, 'The correct renormalization of the gluon operator in a covariant gauge', *Nucl. Phys.* **B379**, 143 (1992).

[75] S. A. Larin, T. van Ritbergen, and J. A. M. Vermaseren, 'The next next-to-leading QCD approximation for nonsinglet moments of deep inelastic structure functions', *Nucl. Phys.* **B427**, 41 (1994).

[76] S. A. Larin, P. Nogueira, T. van Ritbergen, and J. A. M. Vermaseren, 'The 3-loop QCD calculation of the moments of deep inelastic structure functions', *Nucl. Phys.* **B492**, 338 (1997), `hep-ph/9605317`.

[77] A. Retey and J. A. M. Vermaseren, 'Some higher moments of deep inelastic structure functions at next-to-next-to leading order of perturbative QCD', *Nucl. Phys.* **B604**, 281 (2001), `hep-ph/0007294`.

[78] S. Moch, J. A. M. Vermaseren, and A. Vogt, 'The three-loop splitting functions in QCD: The non-singlet case', *Nucl. Phys.* **B688**, 101 (2004), `hep-ph/0403192`.

[79] A. Vogt, S. Moch, and J. A. M. Vermaseren, 'The three-loop splitting functions in QCD: The singlet case', *Nucl. Phys.* **B691**, 129 (2004), `hep-ph/0404111`.

[80] W. Celmaster and D. W. Sivers, 'Studies in the renormalization prescription dependence of perturbative calculations', *Phys. Rev.* **D23**, 227 (1981).

[81] D. J. Gross and C. H. Llewellyn Smith, 'High-energy neutrino - nucleon scattering, current algebra and partons', *Nucl. Phys.* **B14**, 337 (1969).

[82] S. G. Gorishnii and S. A. Larin, 'QCD corrections to the parton model rules for structure functions of deep inelastic scattering', *Phys. Lett.* **B172**, 109 (1986).

[83] E. B. Zijlstra and W. L. van Neerven, 'Order α_s^2 correction to the structure function $F_3(x, Q^2)$ in deep inelastic neutrino - hadron scattering', *Phys. Lett.* **B297**, 377 (1992).

[84] S. A. Larin and J. A. M. Vermaseren, 'The α_s^3 corrections to the bjorken sum rule for polarized electroproduction and to the Gross-Llewellyn Smith sum rule', *Phys. Lett.* **B259**, 345 (1991).

[85] J. D. Bjorken, 'Asymptotic sum rules at infinite momentum', *Phys. Rev.* **179**, 1547 (1969).

[86] G. Altarelli, R. K. Ellis, and G. Martinelli, 'Leptoproduction and drell-yan processes beyond the leading approximation in chromodynamics', *Nucl. Phys.* **B143**, 521 (1978).

[87] S. G. Gorishnii, S. A. Larin, F. V. Tkachov, and K. G. Chetyrkin, 'Higher QCD corrections to the Bjorken sum rule', *Phys. Lett.* **B137**, 230 (1984).

[88] S. A. Larin, F. V. Tkachov, and J. A. M. Vermaseren, 'The α_s^3 correction to the Bjorken sum rule', *Phys. Rev. Lett.* **66**, 862 (1991).

[89] J. D. Bjorken, 'Applications of the chiral U(6) $\otimes$ U(6) algebra of current densities', *Phys. Rev.* **148**, 1467 (1966).

[90] J. D. Bjorken, 'Inelastic scattering of polarized leptons from polarized nucleons', *Phys. Rev.* **D1**, 1376 (1970).

[91] V. M. Braun and A. V. Kolesnichenko, 'Power corrections to Bjorken and Gross-Llewellyn Smith sum rules in QCD', *Nucl. Phys.* **B283**, 723 (1987).

[92] I. I. Balitsky, V. M. Braun, and A. V. Kolesnichenko, 'Power corrections $1/Q^2$ to parton sum rules for deep inelastic scattering from polarized targets', *Phys. Lett.* **B242**, 245 (1990), `hep-ph/9310316`.

[93] 'Erratum', *Phys. Lett.* **B318**, 648 (1993).

[94] S. Fajfer and R. J. Oakes, 'Nonperturbative QCD corrections to the Bjorken and Gross-Llewellyn Smith sum rules', *Phys. Lett.* **B163**, 385 (1985).

[95] U.-K. Yang, A. Bodek, and Q. Fan, 'Parton distributions, d/u, and higher twists at high x', (1998), `hep-ph/9806457`.

[96] M. Dasgupta and B. R. Webber, 'Power corrections and renormalons in deep inelastic structure functions', *Phys. Lett.* **B382**, 273 (1996), `hep-ph/9604388`.

[97] J. H. Kim *et al.*, 'A measurement of $\alpha_s(Q^2)$ from the Gross-Llewellyn Smith sum rule', *Phys. Rev. Lett.* **81**, 3595 (1998), `hep-ex/9808015`.

[98] M. Anselmino, A. Efremov, and E. Leader, 'The theory and phenomenology of polarized deep inelastic scattering', *Phys. Rept.* **261**, 1 (1995), `hep-ph/9501369`.

[99] B. Adeva *et al.*, 'Measurement of the spin dependent structure function $g_1(x)$ of the deuteron', *Phys. Lett.* **B302**, 533 (1993).

[100] B. Adeva *et al.*, 'Combined analysis of world data on nucleon spin structure functions', *Phys. Lett.* **B320**, 400 (1994).

[101] P. L. Anthony *et al.*, 'Determination of the neutron spin structure function.', *Phys. Rev. Lett.* **71**, 959 (1993).

[102] A. Airapetian *et al.*, 'The Q^2-dependence of the generalised Gerasimov-Dell-Hearn integral for the deuteron, proton and neutron', *Eur. Phys. J.* **C26**, 527 (2003), `hep-ex/0210047`.

[103] H. D. Politzer, 'Stevenson's optimized perturbation theory applied to factorization and mass scheme dependence', *Nucl. Phys.* **B194**, 493 (1982).

[104] W. A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta, 'Deep inelastic scattering beyond the leading order in asymptotically free gauge theories', *Phys. Rev.* **D18**, 3998 (1978).

[105] D. W. Duke, J. D. Kimel, and G. A. Sowell, 'Fourth order QCD corrections to the longitudinal coefficient function in deep inelastic scattering', *Phys. Rev.* **D25**, 71 (1982).

[106] A. Devoto, D. W. Duke, J. D. Kimel, and G. A. Sowell, 'Analytic calculation of the fourth order quantum chromodynamic contribution to the nonsinglet quark longitudinal structure function', *Phys. Rev.* **D30**, 541 (1984).

[107] D. I. Kazakov and A. V. Kotikov, 'Total α_s correction to deep inelastic scattering cross-section ratio, $R = \sigma_L/\sigma_t$ in QCD. calculation of longitudinal structure function', *Nucl. Phys.* **B307**, 721 (1988).

[108] D. I. Kazakov and A. V. Kotikov, 'On the value of the α_s correction to the Callan-Gross relation', *Phys. Lett.* **B291**, 171 (1992).

[109] D. I. Kazakov, *et al.*, 'Complete quartic α_s^2 correction to the deep inelastic longitudinal structure function F_L in QCD', *Phys. Rev. Lett.* **65**, 1535 (1990).

[110] W. L. van Neerven and E. B. Zijlstra, 'Order α_s^2 contributions to the deep inelastic wilson coefficient', *Phys. Lett.* **B272**, 127 (1991).

[111] E. B. Zijlstra and W. L. van Neerven, 'Contribution of the second order gluonic Wilson coefficient to the deep inelastic structure function', *Phys. Lett.* **B273**, 476 (1991).

[112] E. B. Zijlstra and W. L. van Neerven, 'Order α_s^2 QCD corrections to the deep inelastic proton structure functions F_2 and F_L', *Nucl. Phys.* **B383**, 525 (1992).

[113] S. Moch and J. A. M. Vermaseren, 'Deep inelastic structure functions at two loops', *Nucl. Phys.* **B573**, 853 (2000), `hep-ph/9912355`.

[114] F. J. Dyson, 'Divergence of perturbation theory in quantum electrodynamics', *Phys. Rev.* **85**, 631 (1952).

[115] M. Beneke, 'Renormalons', *Phys. Rept.* **317**, 1 (1999), `hep-ph/9807443`.

[116] M. Beneke and V. M. Braun, 'Renormalons and power corrections', (2000), `hep-ph/0010208`.

[117] C. M. Bender and T. T. Wu, 'Large order behavior of perturbation theory', *Phys. Rev. Lett.* **27**, 461 (1971).

[118] C. M. Bender and T. T. Wu, 'Statistical analysis of Feynman diagrams', *Phys. Rev. Lett.* **37**, 117 (1976).

[119] G. Parisi, 'Asymptotic estimates of Feynman diagrams', *Phys. Lett.* **B68**, 361 (1977).

[120] I. I. Balitsky, 'Instanton induced asymptotics of perturbative series for $R(e^+e^- \to$ hadrons) and $\Lambda_{\overline{\rm MS}}$ QCD', *Phys. Lett.* **B273**, 282 (1991).

[121] A. Babansky and I. Balitsky, 'Renormalons as dilatation modes in the functional space', *Phys. Rev. Lett.* **85**, 4211 (2000), hep-ph/0008062.

[122] J. Fischer, 'The role of power expansions in quantum field theory', *Int. J. Mod. Phys.* **A12**, 3625 (1997), hep-ph/9704351.

[123] R. J. G. Alves, 'A convergent reformulation of perturbative QCD', Ph.D. thesis, University of Durham (2000).

[124] G. Hardy, *Divergent Series* (Clarendon Press, 1949).

[125] F. J. Dyson, 'The s matrix in quantum electrodynamics', *Phys. Rev.* **75**, 1736 (1949).

[126] J. S. Schwinger, 'On the green's functions of quantized fields. 1', *Proc. Nat. Acad. Sci.* **37**, 452 (1951).

[127] J. S. Schwinger, 'On the green's functions of quantized fields. 2', *Proc. Nat. Acad. Sci.* **37**, 455 (1951).

[128] G. Parisi, 'Singularities of the Borel transform in renormalizable theories', *Phys. Lett.* **B76**, 65 (1978).

[129] G. Parisi, 'The Borel transform and the renormalization group', *Phys. Rept.* **49**, 215 (1979).

[130] M. Neubert, 'Scale setting in QCD and the momentum flow in Feynman diagrams', *Phys. Rev.* **D51**, 5924 (1995), hep-ph/9412265.

[131] S. Peris and E. de Rafael, 'Low-energy QCD and ultraviolet renormalons', *Nucl. Phys.* **B500**, 325 (1997), hep-ph/9701418.

[132] M. Baker and K. Johnson, 'Quantum electrodynamics at small distances', *Phys. Rev.* **183**, 1292 (1969).

[133] H. Kawai, T. Kinoshita, and Y. Okamoto, 'Asymptotic photon propagator and higher order QED Callan-Symanzik beta function', *Phys. Lett.* **B260**, 193 (1991).

[134] D. J. Broadhurst, 'Large N expansion of QED: Asymptotic photon propagator and contributions to the muon anomaly, for any number of loops', *Z. Phys.* **C58**, 339 (1993).

[135] D. J. Broadhurst and A. L. Kataev, 'Connections between deep inelastic and annihilation processes at next to next-to-leading order and beyond', *Phys. Lett.* **B315**, 179 (1993), hep-ph/9308274.

[136] A. Denner, G. Weiglein, and S. Dittmaier, 'Gauge invariance of green functions: Background field method versus pinch technique', *Phys. Lett.* **B333**, 420 (1994), hep-ph/9406204.

[137] J. M. Cornwall, 'Dynamical mass generation in continuum QCD', *Phys. Rev.* **D26**, 1453 (1982).

[138] J. M. Cornwall and J. Papavassiliou, 'Gauge invariant three gluon vertex in QCD', *Phys. Rev.* **D40**, 3474 (1989).

[139] N. J. Watson, 'Universality of the pinch technique gauge boson selfenergies', *Phys. Lett.* **B349**, 155 (1995), hep-ph/9412319.

[140] N. J. Watson, 'The gauge-independent QCD effective charge', *Nucl. Phys.* **B494**, 388 (1997), hep-ph/9606381.

[141] D. Binosi and J. Papavassiliou, 'The pinch technique to all orders', *Phys. Rev.* **D66**, 111901 (2002), hep-ph/0208189.

[142] C. N. Lovett-Turner and C. J. Maxwell, 'Renormalon singularities of the QCD vacuum polarization function to leading order in $1/N_f$.', *Nucl. Phys.* **B432**, 147 (1994), hep-ph/9407268.

[143] C. N. Lovett-Turner and C. J. Maxwell, 'All orders renormalon resummations for some QCD observables', *Nucl. Phys.* **B452**, 188 (1995), hep-ph/9505224.

[144] G. 't Hooft, 'Some observations on quantum chromodynamics', in *Deeper Pathways in High-Energy Physics*: Proceedings of orbis scientiae, 1977, Miami. B. Kursunoglu, A. Perlmutteand and L. F.Scott (eds.).

[145] G. Parisi, 'On infrared divergences', *Nucl. Phys.* **B150**, 163 (1979).

[146] A. H. Mueller, 'The QCD perturbation series', In *QCD: 20 years later*: Proceedings of Workshop in Aachen, Germany, June 9-13, 1992, P. M. Zerwas H. A. and Kastrup (eds.). World Scientific, 1993.

[147] G. Grunberg, 'Perturbation theory and condensates', *Phys. Lett.* **B325**, 441 (1994).

[148] A. H. Mueller, 'On the structure of infrared renormalons in physical processes at high-energies', *Nucl. Phys.* **B250**, 327 (1985).

[149] F. David, 'On the ambiguity of composite operators, I.R. renormalons and the status of the operator product expansion', *Nucl. Phys.* **B234**, 237 (1984).

[150] F. David, 'The operator product expansion and renormalons: A comment', *Nucl. Phys.* **B263**, 637 (1986).

[151] E. Gardi and G. Grunberg, 'Power corrections in the single dressed gluon approximation: The average thrust as a case study', *JHEP* **11**, 016 (1999), hep-ph/9908458.

[152] P. M. Brooks and C. J. Maxwell, 'Infrared freezing of Euclidean QCD observables', *Phys. Rev.* **D74**, 065012 (2006), hep-ph/0604267.

[153] M. Beneke, 'Large order perturbation theory for a physical quantity', *Nucl. Phys.* **B405**, 424 (1993).

[154] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, 1965).

[155] D. J. Broadhurst and A. G. Grozin, 'Matching QCD and HQET heavy - light currents at two loops and beyond', *Phys. Rev.* **D52**, 4082 (1995), hep-ph/9410240.

[156] M. Beneke and V. M. Braun, 'Naive non-Abelianization and resummation of fermion bubble chains', *Phys. Lett.* **B348**, 513 (1995), hep-ph/9411229.

[157] S. J. Brodsky, E. Gardi, G. Grunberg, and J. Rathsman, 'Disentangling running coupling and conformal effects in QCD', *Phys. Rev.* **D63**, 094017 (2001), hep-ph/0002065.

[158] A. V. Nesterenko and J. Papavassiliou, 'A novel integral representation for the Adler function', *J. Phys.* **G32**, 1025 (2006), hep-ph/0511215.

[159] D. J. Broadhurst and A. L. Kataev, 'Bjorken unpolarized and polarized sum rules: Comparative analysis of large-N_f expansions', *Phys. Lett.* **B544**, 154 (2002), hep-ph/0207261.

[160] M. Beneke, V. M. Braun, and N. Kivel, 'The operator product expansion, non-perturbative couplings and the Landau pole: Lessons from the $O(N)$ sigma-model', *Phys. Lett.* **B443**, 308 (1998), hep-ph/9809287.

[161] I. Caprini and J. Fischer, 'On the infrared freezing of perturbative QCD in the Minkowskian region', *Phys. Rev.* **D71**, 094017 (2005), hep-ph/0505016.

[162] D. V. Shirkov and I. L. Solovtsov, 'Analytic model for the QCD running coupling with universal $\bar{\alpha}_s(0)$ value', *Phys. Rev. Lett.* **79**, 1209 (1997), hep-ph/9704333.

[163] D. V. Shirkov, 'Analytic perturbation theory in analyzing some QCD observables', *Eur. Phys. J.* **C22**, 331 (2001), hep-ph/0107282.

[164] A. M. Badalian and Y. A. Simonov, 'Freezing of $\alpha_s(Q^2)$ in e^+e^- annihilation', *Phys. Atom. Nucl.* **60**, 630 (1997).

[165] S. G. Gorishnii, A. L. Kataev, and S. A. Larin, 'Three loop corrections of order $\mathcal{O}(M^2)$ to the correlator of electromagnetic quark currents', *Nuovo Cim.* **A92**, 119 (1986).

[166] E. C. Poggio, H. R. Quinn, and S. Weinberg, 'Smearing the quark model', *Phys. Rev.* **D13**, 1958 (1976).

[167] M. Abramowitz and I. A. Stegan (eds.), *Handbook of Mathematical Functions* (Dover, 1964).

[168] I. Caprini and M. Neubert, 'Borel summation and momentum-plane analyticity in perturbative QCD', *JHEP* **03**, 007 (1999), hep-ph/9902244.

[169] J. Chyla, A. Kataev, and S. Larin, 'Renormalization scheme dependence and infrared behavior in e^+e^- annihilations and tau lepton decay at the next-to - next-to-leading order of perturbative QCD', *Phys. Lett.* **B267**, 269 (1991).

[170] C. J. Maxwell and D. G. Tonge, 'RS-invariant all-orders renormalon resummations for some QCD observables', *Nucl. Phys.* **B481**, 681 (1996), hep-ph/9606392.

[171] C. J. Maxwell and D. G. Tonge, 'The uncertainty in $\alpha_s(M_Z^2)$ determined from hadronic tau decay measurements', *Nucl. Phys.* **B535**, 19 (1998), hep-ph/9705314.

[172] J. Chyla and A. L. Kataev, 'Next-to-next-to-leading order QCD analysis of the Gross- Llewellyn-Smith sum rule and the higher twist effects', *Phys. Lett.* **B297**, 385 (1992), hep-ph/9209213.

[173] J. R. Ellis, E. Gardi, M. Karliner, and M. A. Samuel, 'Pade approximants, Borel transforms and renormalons: the Bjorken sum rule as a case study', *Phys. Lett.* **B366**, 268 (1996), hep-ph/9509312.

[174] G. Altarelli, R. D. Ball, S. Forte, and G. Ridolfi, 'Determination of the Bjorken sum and strong coupling from polarized structure functions', *Nucl. Phys.* **B496**, 337 (1997), hep-ph/9701289.

[175] T. Lee, 'Renormalons beyond one-loop', *Phys. Rev.* **D56**, 1091 (1997), hep-th/9611010.

[176] T. Lee, 'Normalization constants of large order behavior', *Phys. Lett.* **B462**, 1 (1999), hep-ph/9908225.

[177] K. A. Milton, I. L. Solovtsov, and O. P. Solovtsova, 'The Gross-Llewellyn Smith sum rule in the analytic approach to perturbative QCD', *Phys. Rev.* **D60**, 016001 (1999), hep-ph/9809513.

[178] C. Contreras, G. Cvetic, K. S. Jeong, and T. Lee, 'Extraction of α_s from the Gross-Llewellyn Smith sum rule using Borel resummation', *Phys. Rev.* **D66**, 054006 (2002), hep-ph/0203201.

[179] F. Campanario and A. Pineda, 'Fit to the Bjorken, Ellis-Jaffe and Gross-Llewellyn-Smith sum rules in a renormalon based approach', *Phys. Rev.* **D72**, 056008 (2005), hep-ph/0508217.

[180] C. J. Maxwell and A. Mirjalili, 'Renormalon-inspired resummations for vector and scalar correlators: Estimating the uncertainty in $\alpha_s(M_\tau^2)$ and $\alpha_s(M_Z^2)$', *Nucl. Phys.* **B611**, 423 (2001), hep-ph/0103164.

[181] K. Abe *et al.*, 'Measurements of the proton and deuteron spin structure functions g_1 and g_2', *Phys. Rev.* **D58**, 112003 (1998), hep-ph/9802357.

[182] B. Adeva *et al.*, 'A next-to-leading order QCD analysis of the spin structure function g_1', *Phys. Rev.* **D58**, 112002 (1998).

[183] P. L. Anthony *et al.*, 'Measurements of the Q^2 dependence of the proton and neutron spin structure functions g_1^p and g_1^n', *Phys. Lett.* **B493**, 19 (2000), hep-ph/0007248.

[184] A. Deur *et al.*, 'Experimental determination of the evolution of the Bjorken integral at low Q^2', *Phys. Rev. Lett.* **93**, 212001 (2004), hep-ex/0407007.

[185] P. M. Brooks and C. J. Maxwell, 'Comparison of NNLO and all-orders estimates of corrections to the GLS and Bjorken sum rules', (2006), hep-ph/0608339.

[186] S. J. Burby and C. J. Maxwell, 'Direct extraction of QCD $\Lambda_{\overline{\rm MS}}$ from e^+e^- jet observables', *Nucl. Phys.* **B609**, 193 (2001), hep-ph/0011203.

[187] W. Bernreuther and W. Wetzel, 'Decoupling of heavy quarks in the minimal subtraction scheme', *Nucl. Phys.* **B197**, 228 (1982).

[188] T. Appelquist and J. Carazzone, 'Infrared singularities and massive fields', *Phys. Rev.* **D11**, 2856 (1975).

[189] S. A. Larin, T. van Ritbergen, and J. A. M. Vermaseren, 'The large quark mass expansion of $\Gamma(Z_0 \to \text{hadrons})$ and $\Gamma(\tau \to \nu_\tau + \text{hadrons})$ in the order α_s^3', *Nucl. Phys.* **B438**, 278 (1995), hep-ph/9411260.

[190] S. I. Alekhin and A. L. Kataev, 'QCD effects to Bjorken unpolarized sum rule for νN deep- inelastic scattering', *J. Phys.* **G29**, 1993 (2003), hep-ph/0209165.

[191] A. L. Kataev, 'Renormalons at the boundaries between perturbative and non-perturbative QCD', *Mod. Phys. Lett.* **A20**, 2007 (2005), hep-ph/0505230.

[192] J. Santiago and F. J. Yndurain, 'Calculation of electroproduction to NNLO and precision determination of α_s', *Nucl. Phys.* **B563**, 45 (1999), hep-ph/9904344.

[193] J. Santiago and F. J. Yndurain, 'Improved calculation of F_2 in electroproduction and xF_3 in neutrino scattering to NNLO and determination of α_s', *Nucl. Phys.* **B611**, 447 (2001), hep-ph/0102247.

[194] C. J. Maxwell and A. Mirjalili, 'Direct extraction of QCD $\Lambda_{\overline{\text{MS}}}$ from moments of structure functions in neutrino nucleon scattering, using the CORGI approach', *Nucl. Phys.* **B645**, 298 (2002), hep-ph/0207069.

[195] A. Mirjalili, 'Scale dependence and renormalon inspired resummations for some QCD observables', Ph.D. thesis, University of Durham (2001).

[196] P. M. Brooks and C. J. Maxwell, 'Improved analysis of moments of F_3 in neutrino nucleon scattering using the Bernstein polynomial method', (2006), hep-ph/0610137.

[197] P. M. Stevenson and H. D. Politzer, 'Optimized perturbation theory applied to factorization scheme dependence', *Nucl. Phys.* **B277**, 758 (1986).

[198] H. Georgi and H. D. Politzer, 'Freedom at moderate energies: Masses in color dynamics', *Phys. Rev.* **D14**, 1829 (1976).

[199] O. Nachtmann, 'Positivity constraints for anomalous dimensions', *Nucl. Phys.* **B63**, 237 (1973).

[200] W. G. Seligman *et al.*, 'Improved determination of α_s from neutrino nucleon scattering', *Phys. Rev. Lett.* **79**, 1213 (1997).

[201] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, 'Parton distributions incorporating QED contributions', *Eur. Phys. J.* **C39**, 155 (2005), hep-ph/0411040.

[202] J. Pumplin *et al.*, 'New generation of parton distributions with uncertainties from global QCD analysis', *JHEP* **07**, 012 (2002), hep-ph/0201195.

[203] S. Bernstein, 'Démonstration du théorème de weierstrass fondée sur le calcul des probabilities.', *Comm. Soc. Math. Kharkov* **13**, 1 (1912).

[204] K. Adel, F. Barreiro, and F. J. Yndurain, 'Theory of small x deep inelastic scattering: NLO evaluations, and low Q^2 analysis', *Nucl. Phys.* **B495**, 221 (1997), hep-ph/9610380.

[205] R. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences* (John Wiley and Sons Ltd, 1996).

[206] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, 'NNLO global parton analysis', *Phys. Lett.* **B531**, 216 (2002), hep-ph/0201127.

[207] A. J. Buras, 'Asymptotic freedom in deep inelastic processes in the leading order and beyond', *Rev. Mod. Phys.* **52**, 199 (1980).

[208] P. M. Brooks and C. J. Maxwell, 'In preparation', (2006).

[209] L. W. Whitlow, *et al.*, 'Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross-sections', *Phys. Lett.* **B282**, 475 (1992).

[210] A. C. Benvenuti *et al.*, 'A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2', *Phys. Lett.* **B223**, 485 (1989).

[211] M. R. Adams *et al.*, 'Proton and deuteron structure functions in muon scattering at 470-GeV', *Phys. Rev.* **D54**, 3006 (1996).

[212] M. Derrick *et al.*, 'Measurement of the F_2 structure function in deep inelastic $e + p$ scattering using 1994 data from the ZEUS detector at HERA', *Z. Phys.* **C72**, 399 (1996), hep-ex/9607002.

[213] S. Aid *et al.*, 'A measurement and QCD analysis of the proton structure function $F_2(x, Q^2)$ at HERA', *Nucl. Phys.* **B470**, 3 (1996), hep-ex/9603004.

[214] C. Adloff *et al.*, 'Measurement of neutral and charged current cross sections in electron proton collisions at high Q^2', *Eur. Phys. J.* **C19**, 269 (2001), hep-ex/0012052.

[215] M. Arneodo *et al.*, 'Measurement of the proton and deuteron structure functions, F_2^p and F_2^d, and of the ratio σ_L/σ_T', *Nucl. Phys.* **B483**, 3 (1997), hep-ph/9610231.